

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 3.Sep.03	3. REPORT TYPE AND DATES COVERED THESIS		
4. TITLE AND SUBTITLE "CHARACTERIZATION OF STREPTOCOCCUS SANGUIS MUTANTS GENERATED BY SIGNATURE TAGGED MUTAGENESIS"		5. FUNDING NUMBERS		
6. AUTHOR(S) MAJ NOE JODY C				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) VIRGINIA COMMONWEALTH UNIV MED COL OF VA		8. PERFORMING ORGANIZATION REPORT NUMBER  CI02-1276		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) THE DEPARTMENT OF THE AIR FORCE AFIT/CIA, BLDG 125 2950 P STREET WPAFB OH 45433		10. SPONSORING/MONITORING AGENCY REPORT NUMBER		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION AVAILABILITY STATEMENT Unlimited distribution In Accordance With AFI 35-205/AFIT Sup 1		12b. DISTRIBUTION CODE		
13. ABSTRACT (Maximum 200 words)				
<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;"> <b>DISTRIBUTION STATEMENT A</b>            Approved for Public Release            Distribution Unlimited         </div> <div style="font-size: 2em; font-weight: bold;">20030916 009</div> </div>				
14. SUBJECT TERMS		15. NUMBER OF PAGES 76		
		16. PRICE CODE		
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT	

**CHARACTERIZATION OF *STREPTOCOCCUS SANGUIS* MUTANTS  
GENERATED BY SIGNATURE-TAGGED MUTAGENESIS**

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

JODY CHRISTINE NOE

A.A.S. 1980, Kellogg Community College  
B.S. 1982, Eastern Michigan University

Director: Dr Todd Kitten  
Assistant Professor, Department Of Microbiology and Immunology and  
The Philips Institute of Oral and Craniofacial Molecular Biology

Virginia Commonwealth University  
Richmond, Virginia  
AUGUST, 2003

THE VIEWS EXPRESSED IN THIS ARTICLE ARE THOSE OF  
THE AUTHOR AND DO NOT REFLECT THE OFFICIAL  
POLICY OR POSITION OF THE UNITED STATES AIR  
FORCE, DEPARTMENT OF DEFENSE, OR THE U.S.  
GOVERNMENT

### Acknowledgement

I would like to thank several individuals. The most important of whom are my husband . . . , whose support and love have made this journey possible and my children, . . . who have been understanding and made numerous sacrifices so that I could have this and many other career opportunities. I am truly blessed. I would also like to thank Sehmi Paik and Amanda DeRocco for their friendship and academic help. I'd like to thank all the members of my committee but especially Dr. Todd Kitten, who as my advisor surely deserves a medal for his unflagging patience and forbearance in his dealings with me and Dr Gail Christie who was my teacher, my tutor and my friend. Her support during a critical period in my life will never be forgotten. Last of all, I'd like to thank the United States Air Force Institute of Technology for the scholarship that allowed me the opportunity to earn this degree.

## Table of Contents

	Page
Acknowledgements .....	ii
List of Tables .....	v
List of Figures .....	vi
Abstract .....	vii
Chapter	
1 Introduction .....	1
Infective Endocarditis .....	1
<i>Streptococcus sanguis</i> .....	2
Treatment of Endocarditis .....	2
Signature Tagged Mutagenesis .....	4
First STM Experiment .....	11
2 Materials and Methods .....	14
Bacterial Strains, Plasmids and Media .....	14
Chromosomal DNA Isolation and Quantification .....	15
Mutant Characterization by Cloning .....	17
Mutant Characterization by AP-PCR .....	22
Polymerase Chain Reaction .....	25
DNA Gel Electrophoresis and Southern Blotting .....	25
DNA Sequence Analysis .....	26
3 Results .....	27
Growth Studies .....	27
DNA Isolation and Quantification .....	30
Amplification of Signature Tags .....	32
Southern Blot Results .....	35
Characterization of Mutant 1-39 .....	38
Characterization of Mutant 1-17 .....	53

Characterization of Mutants 1-36 and 1-37 .....	60
4 Discussion .....	64
Literature Cited.....	72

List of Tables

Table	Page
1. Primers used in this study.....	20
2. Primer Pairs used for AP-PCR. ....	23
3. DNA Quantification Results .....	31

# List of Figures

Figure	Page
1. pJFP1 with Mini-Transposon .....	6
2. Signature-Tagged Mutagenesis .....	10
3. Dot Blot Analysis .....	13
4. Growth Studies. ....	29
5. Signature-Tag Amplification.....	34
6. Southern Blots of Mutant Strains .....	37
7. 1-39 Cloning Experiment .....	40
8. Position of Sequencing Primers used in Cloning, 1-39 .....	43
9. Principle of AP-PCR .....	46
10. Strategy View of Mutant 1-39 .....	50
11. Mutant 1-39, Southern Blot.....	52
12. Strategy View of Mutant 1-17 .....	56
13. Mutant 1-17 Southern Blot.....	59
14. Strategy View of Mutants 1-36 and 1-37 .....	63
15. Hypothetical Model of Mutant 1-39 Generation .....	70



## Abstract

### CHARACTERIZATION OF *STREPTOCOCCUS SANGUIS* MUTANTS GENERATED BY SIGNATURE-TAGGED MUTAGENESIS

By **Jody Christine Noe, M.S**

A Thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2003

Major Director: Dr Todd Kitten  
Assistant Professor, Department Of Microbiology And Immunology and  
The Philips Institute of Oral and Craniofacial Molecular Biology

*Streptococcus sanguis* belongs to the viridans group of oral streptococci, which cause 40-50% of human native valve endocarditis. Identification of new virulence factors is important for drug and vaccine development. A transposon-based random mutagenesis technique called signature-tagged mutagenesis (STM) was used to search for new virulence factors. Four signature-tagged mutants exhibiting decreased virulence in a rabbit model were chosen for characterization. Growth studies suggested that poor growth may have contributed to their decreased virulence. Next, mutant DNA was cloned into *Escherichia coli* to characterize the sequences flanking the transposon. Lastly, all mutants were characterized using arbitrarily primed polymerase chain reaction.

It was determined that plasmid DNA was inserted with the transposon in three of the four mutants. Additionally, in two of the mutants, the transposon was flanked by apparently non-contiguous sequences. These surprising results may be due to unforeseen events occurring during the *in vitro* transposition stage of STM.

## Introduction

### Infective Endocarditis

Infective endocarditis (IE) is a life threatening disease that affects approximately 22,000 people in the United States each year. On average 2000 of those patients will die (Vlessis, Hovaguimian et al. 1996; Dyson, Barnes et al. 1999; Kurland, Enghoff et al. 1999). Infective endocarditis is caused by a bacterial infection in the endocardium of the heart (Roberts, Krieger et al. 1979; Douglas, Heath et al. 1993; Vlessis, Hovaguimian et al. 1996; Dyson, Barnes et al. 1999; Kurland, Enghoff et al. 1999). Typically one of the four heart valves is the focus of infection; however, it can also occur at the site of a septal defect such as on the chordae tendineae, or on the mural endocardium (Mylonakis and Calderwood 2001).

Patients at risk for IE are those who have one or more of the following: mitral (bicuspid) valve prolapse; prosthetic heart valve(s); acquired valvular dysfunction (such as that caused by rheumatic heart disease); or congenital conditions that lead to turbulent blood flow through the heart valves.

Infective endocarditis can be divided into two main types, native valve endocarditis (NVE) and prosthetic valve endocarditis (PVE). PVE occurs when the infection involves an artificial heart valve and is most typically caused by *Staphylococcus*

*aureus* (Cabell, Jollis et al. 2002). On the other hand, NVE is an infection that normally occurs in a previously damaged native heart valve and is more often caused by one of the viridans streptococci.

### *Streptococcus sanguis*

The viridans streptococci are gram positive cocci that grow in chains and exhibit  $\alpha$  hemolysis when grown on sheep blood agar. This hemolysis appears as a green sheen on the agar. Many species make up the viridans streptococci. These include, but are not limited to *S. sanguis*, *S. mutans*, *S. bovis*, *S. salivarius*, *S. pneumoniae* and *S. anginosus*. However, of all the viridans streptococci, the one implicated most often in infective endocarditis is *Streptococcus sanguis* (Roberts, Krieger et al. 1979; Douglas, Heath et al. 1993; Vlessis, Hovaguimian et al. 1996; Dyson, Barnes et al. 1999; Kurland, Enghoff et al. 1999).

*S. sanguis* is a part of the normal flora of the oral cavity. It has not been shown to cause dental caries but it has been implicated in other, extra-oral diseases including bacteremia in neutropenic patients in addition to infective endocarditis. Also, *S. sanguis* is naturally competent for transformation.

### Treatment of Endocarditis

Treatment of endocarditis is extremely problematic. There is a high incidence of morbidity as well as mortality with this disease (Bayer, Bolger et al. 1998). This is due in part to the nature of the infection. The effectiveness of antibiotics is limited due to formation of a vegetation on the surface of the heart valve. This vegetation is composed of microorganisms, fibrin, platelets and possibly other matrix proteins such as fibronectin

and collagen (Durack 1975; Switalski, Murchison et al. 1987; Moreillon, Overholser et al. 1988; Tart and van de Rijn 1991). The vegetation limits access of antibiotics to the microorganisms thus making it hard to treat affected patients. Therefore in many cases surgical intervention in the form of heart valve replacement is necessary (Vlessis, Hovaguimian et al. 1996).

Microorganisms, and in particular viridans streptococci, can be introduced into the blood stream in a transient manner under a variety of circumstances such as chewing food, brushing and flossing of teeth or having a dental procedure performed (Guntheroth 1984; Roberts, Holzel et al. 1997). Therefore patients with a known heart defect are typically given prophylactic antibiotics prior to having dental procedures performed (Dajani 1998). However, since this provides only short-term protection from the risk of endocarditis, it would be better to develop a vaccine against the viridans streptococci for this at-risk population that would offer continuous protection (Baddour 1999).

In order to develop an effective vaccine for native valve endocarditis it is important to target a protein that is not only surface-exposed but that also has homologs in all the viridans streptococci that have been implicated in the disease (Kitten, Munro et al. 2002). If the vaccine could also protect against *S. aureus*, which is typically found in prosthetic valve endocarditis, it would be even better. However, this is less likely to occur. In the past, vaccine candidates have often been targeted to known virulence factors. Studying the processes needed to successfully establish infective endocarditis allowed for the discovery of many of these virulence factors. These processes include host immune evasion, adherence to the vegetation, and growth of the vegetation (Sullam,

Frank et al. 1993). One problem with this approach is that it doesn't allow for the discovery of virulence genes that we don't know to look for. It is probable that there are many genes important to virulence whose function is not readily recognizable as being important to virulence.

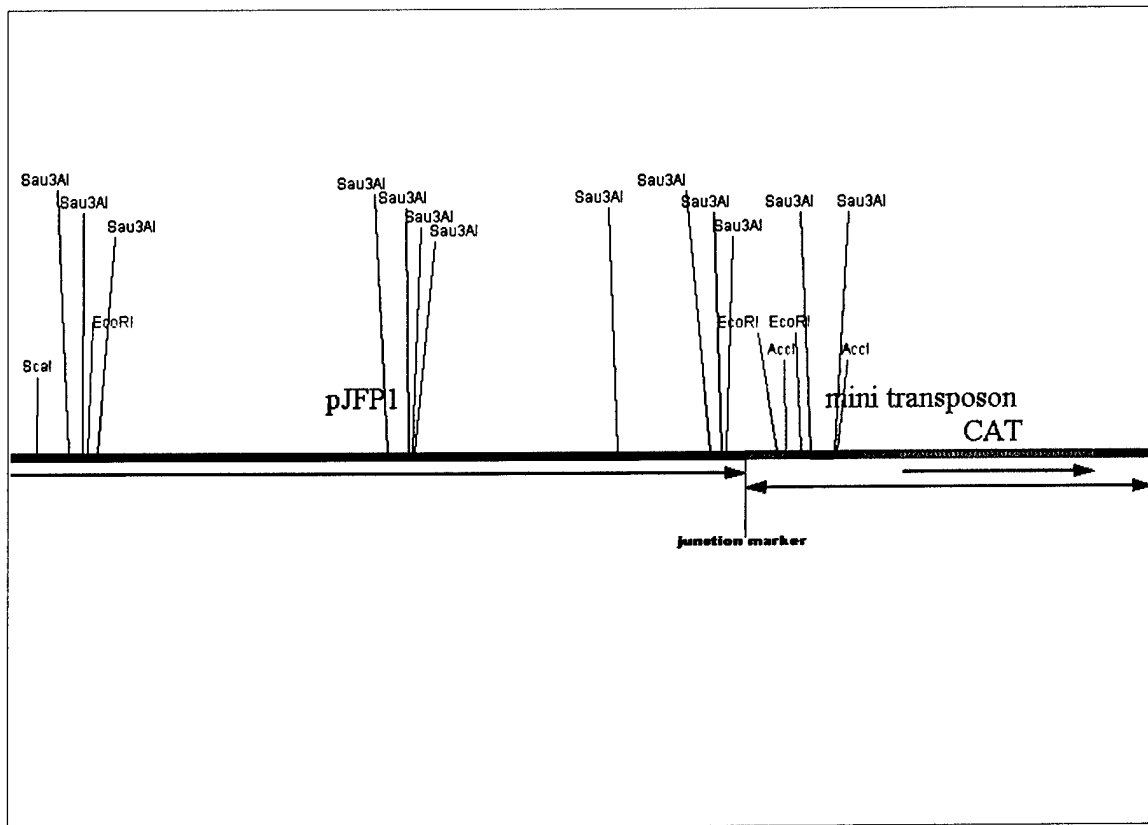
In order to identify these virulence genes, a different approach to the problem is necessary. One way of identifying these genes is by a technique called signature tagged mutagenesis.

### Signature Tagged Mutagenesis

Signature tagged mutagenesis (STM) is a recently developed technique used to identify virulence genes of various microbial pathogens (Unsworth and Holden 2000). In this case, a strain of *S. sanguis* is mutated in such a way that each mutated strain is "labeled" with a unique signature tag by use of transposon mutagenesis (Lampe, Churchill et al. 1996; Akerley, Rubin et al. 1998; Lampe, Grant et al. 1998; Lampe, Akerley et al. 1999). A mini-transposon, from the *mariner* family in plasmid pJFP1, was used to insert uniquely labeled mutations into the DNA of *S. sanguis* strain SK36. A schematic of the transposon and plasmid can be seen in Figure 1.

Figure 1. pJFP1 with Mini-Transposon.

The figure represents a schematic of the linearized pJFP1 plasmid with the mini-transposon inserted. The relative position of the chloramphenicol acetyl transferase gene within the plasmid is depicted. Also shown are the restriction enzyme cut sites for the enzymes used in the Southern blot experiments.





The transposon also carries the chloramphenicol acetyl-transferase (*cat*) gene, which is responsible for conferring resistance to chloramphenicol. The *cat* gene makes it possible to select for organisms that have successfully inserted the mini-transposon into the *S. sanguis* genome. Since the mini-transposon does not encode its own transposase enzyme, the latter is added separately during the *in vitro* transposition step of the protocol. During this step the mini-transposon in plasmid pJFP1, buffer, chromosomal DNA used as a target for transposition, and the transposase are incubated together to allow the transposition reaction to occur. The reaction is then stopped by heating. T4 polymerase and T4 ligase are then added to close the gaps created by the transposition process. Next the mutated DNA fragments are transformed into *S. sanguis* where they are incorporated into the chromosome by homologous recombination.

Dot blots are then made with the plasmid DNA used to create each mutant (Figure 2). The mutated strains are then grown in individual cultures, pooled, and used for two separate steps. The first is to prepare input pool DNA for PCR amplification and signature tag labeling. The second is for inoculating the animal model. We originally used rats but have since found better success using rabbits (Kitten, Munro et al. 2002).

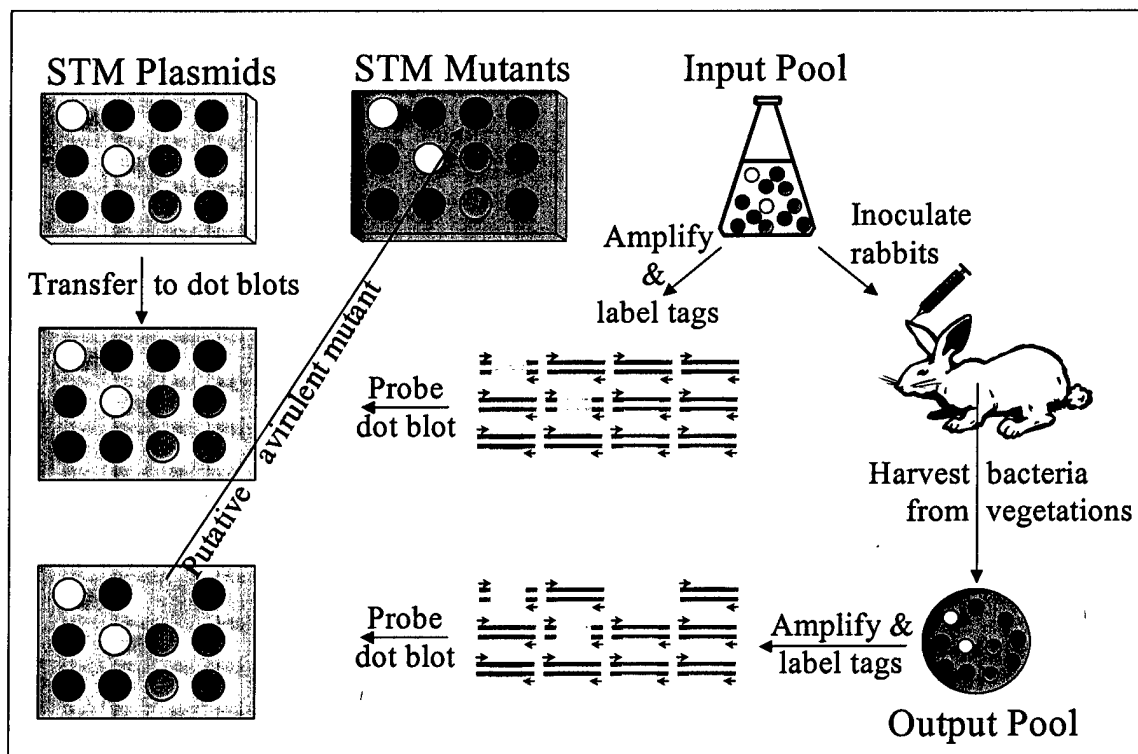
In order to determine virulence factors that are likely to be important in endocarditis in humans, it is necessary to mimic the actual disease state as closely as possible. This is done by traumatizing the heart valve of the rabbits by inserting a catheter through the right carotid artery then through the aortic valve and into the left ventricle. The catheter is kept in this position for the duration of the experiment (Durack,

Beeson et al. 1973). On the second post-operative day the rabbits are infected with a pool of forty different signature-tagged mutant strains of *S. sanguis* (the input pool). One to two days after infection the rabbits are euthanized and the heart valves are recovered.

The mutants that successfully infected the heart valves are recovered from them (the output pool). Next the DNA is extracted from both the input and output pools. The DNA is amplified and the signature tags are labeled. Both probes are hybridized with duplicate dot blots. In this way, we can see which strains were present in the input pool but absent in the output pool. The input pool probe should hybridize to every spot; however, decreased signals in the output pool could signify mutants with decreased virulence and would be further investigated.

Figure 2. Signature-Tagged Mutagenesis.

Dot blots are made from the different STM plasmids, each with a unique signature tag, represented here by a different color. The different strains are then pooled together and divided into two aliquots. One aliquot is set aside for later creation of the input pool probe. The second aliquot of pooled organisms is injected into a rabbit with a previously traumatized heart valve. Later the organisms are recovered from the heart valve (output pool). Avirulent mutants won't survive in the animal. DNA from both the input and output pools is extracted and the signature tags are amplified and used as probes against the previously made dot blots. A change in signal strength could be due to a change in virulence and would be studied further. In the example shown, the blue mutant is present in the input pool, but missing from the output pool, suggesting that it is avirulent.

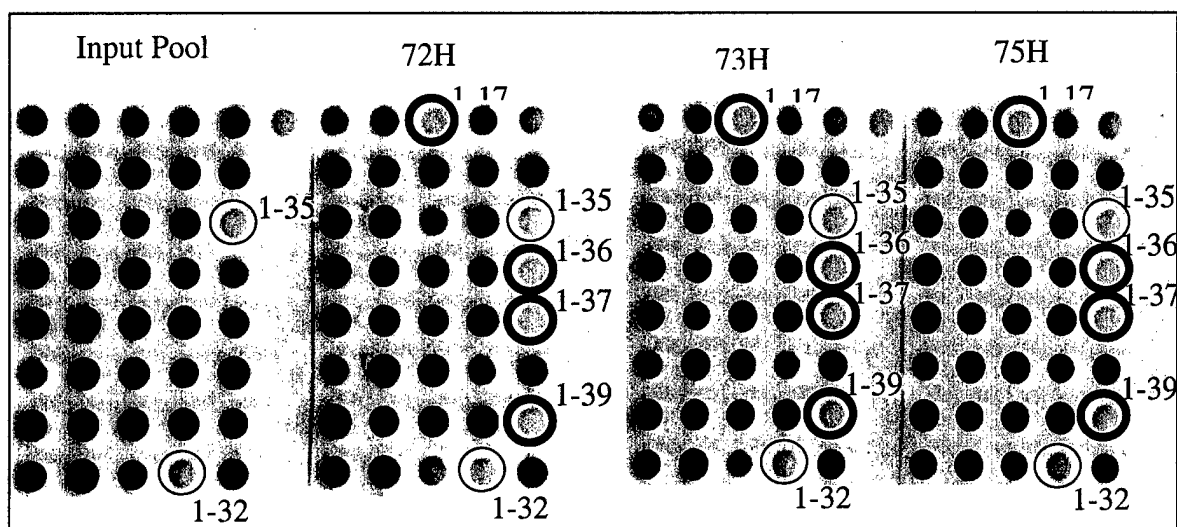


### First STM Experiment

The first pool was made of 40 mutant strains. The input pool had two strains with decreased signals, 1-32 and 1-35 (light circles in Fig. 3). The output pool had six strains with decreased signals – the two from the input pool and four additional strains 1-17, 1-36, 1-37 and 1-39 (dark circles). These putative avirulent strains (1-17, 1-36, 1-37 and 1-39) were chosen for further characterization.

Figure 3. Dot Blot Analysis.

The blots indicated signal variation between the input pool and the output pools. 72H, 73H and 75H indicate three separate animal output pools all showing consistent results. Light circles indicate mutants that had decreased signal in every pool. Dark circles are those that were decreased only in the output pools.



## Materials and Methods

### Bacterial Strains, Plasmids and Media

*Streptococcus sanguis* strain SK36 was obtained from Mogens Kilian at the University of Aarhus in Denmark (Hsu, Cisar et al. 1994). This was the parent strain for the signature-tagged mutagenesis protocol. Selected mutant strains were grown anaerobically at 37 °C in Brain Heart Infusion broth (BHI; Becton Dickinson, Sparks, MD). These were grown both in the presence and the absence of 5µg/ml of chloramphenicol to determine growth rates for the different mutants. Mutant strains included in the growth study were 1-17, 1-36, 1-37 and 1-39 (putative avirulent strains), 1-32 and 1-35 (decreased signal on the initial inoculum) and 1-5 and 1-15 (used as normal controls).

Plasmid pJFP1 was provided by Todd Kitten (Virginia Commonwealth University) (Unpublished data). It was constructed from a parent plasmid, pEMCAT, provided by Andrew Camilli, (Tufts University) and contains a mini-transposon from the *mariner* family (Hava and Camilli 2002). The mini-transposon contains the chloramphenicol acetyl transferase (*cat*) gene, which confers resistance to chloramphenicol.

Plasmid pVA2606 was used for cloning experiments. It was created by Todd



Kitten (Virginia Commonwealth University) by digesting pUC19 with DraI to remove the *bla* gene, which was replaced with an end-filled BamHI fragment from pVA2592 containing the *aphA-3* gene, encoding resistance to kanamycin.

*Escherichia coli* DH10B electrocompetent cells for cloning experiments were obtained from Invitrogen, Carlsbad, CA. Electroporated *E. coli* cells were grown on Luria agar (L agar) made from Luria Broth (LB Broth) added to Bacto agar in amounts specified by the manufacturer (both reagents are from Becton Dickinson, Sparks, MD). The L agar was supplemented by the addition of 50µg/ml of kanamycin, 60 µg/ml of 5-bromo-4-chloro-3-indolyl-β-D-galactoside (x-gal) and 5µg/ml chloramphenicol to select for the desired transformants. Terrific broth (Sambrook and Russell 2001) was supplemented with the addition of 50µg/ml of kanamycin and 5µg/ml chloramphenicol and was also used in the cloning experiments.

#### Chromosomal DNA Isolation and Quantification

Chromosomal DNA isolation and characterization of mutant strains 1-17, 1-32, 1-35, 1-36, 1-37 and 1-39 was performed as described (Kitten, Munro et al. 2000). Strains were grown anaerobically overnight at 37 °C in 3ml of fresh BHI supplemented with 20 mM DL-threonine (BHIT - Fisher Scientific, Pittsburgh, PA) and 5µg/ml chloramphenicol. In the morning an additional 9ml of BHIT was added and the cells were grown for another hour. At that point 0.6 g of glycine (Sigma/Aldrich, St. Louis, MO) was added and the cells were incubated an additional 45 minutes. Cells were harvested by centrifugation (10 min., 3,300 x g , 4 °C).

The resulting pellet was washed twice in 1 ml of cold water. Washed cells were suspended in 0.36 ml GET buffer. GET buffer contains 25% glucose (Fisher Scientific, Pittsburgh, PA) in TE buffer (Tris base, 10 mM, pH 7.4, EDTA 1 mM). Fifty  $\mu$ g of RNase A (Promega, Madison, WI) and 1 mg lysozyme (ICN Biomedicals, Inc., Aurora, OH) were added and the solution incubated for 37  $^{\circ}$ C for 30 minutes. Next 100  $\mu$ g Proteinase K (Invitrogen, Carlsbad, CA) and 54  $\mu$ l of 10% Sodium N-lauroyl sarcosinate (Sarkosyl) was added and the mixture inverted and then incubated for another hour at 37  $^{\circ}$ C or until the solution was clear.

Four hundred microliters of phenol: chloroform: isoamyl alcohol solution (25:24:1; Fisher Scientific, Pittsburgh, PA) were added and the mixture was vigorously mixed and then transferred to a phase lock gel (heavy) tube (Eppendorf/Brinkman, Westbury, NY). The phase lock gel tubes were then centrifuged in a microfuge (12,500 x g, 5 min., room temperature). The aqueous phase was removed and 0.4 ml of  $\text{CHCl}_3$  (Chloroform, Fisher Scientific, Pittsburgh, PA) was added. This mixture was again centrifuged (12,500 g, 5 min, room temperature). The aqueous phase resulting from this step was then saved and the DNA precipitated (Maniatis, Fritsch et al. 1982) as follows. One-tenth volume of 3 M sodium acetate (Fisher Scientific, Pittsburgh, PA) was added and the solution was mixed gently. Next 2.5 volumes or 1 ml of EtOH (whichever was the smaller amount) was added. At this stage the DNA was visible and the tube was centrifuged (12,500 x g, 1 min., room temperature).

The pellet was then rinsed with 70% ethanol and dried in a speedvac system (Automatic Environmental SpeedVac ® System, Thermo Savant, Holbrook, NY). The DNA was then resuspended in 200 µl of TE.

The resulting DNA was quantified by spectrophotometry. Readings for each DNA preparation were obtained at both OD<sub>260</sub> and OD<sub>280</sub>. The OD<sub>260</sub> value was used to calculate the concentration of the DNA sample. The OD<sub>260</sub>:OD<sub>280</sub> ratio was used to determine the purity of the DNA preparation.

#### Mutant Characterization by Cloning

Genomic DNA from mutant 1-39 was partially digested using serial dilutions of the restriction enzyme *Sau3AI* (New England Biolabs, Beverly, MA) according to the manufacturer's directions. Aliquots of the serial diluted DNA fragments were separated by electrophoresis and those reactions containing bands of mainly 2 Kb or larger in size were then ligated to vector plasmid pVA2606, which had been previously digested with *BamHI* (New England Biolabs, Beverly, MA) according to manufacturer's directions. T4 Ligase (Invitrogen, Carlsbad, CA) was used to ligate the vector DNA to the insert (1-39) DNA following the manufacturer's directions.

Ligated DNA was dialyzed by placing 14 µl of the DNA into a 3500 MWCO Slide-A-Lyzer ® Mini dialysis unit (Pierce, Rockford, IL). This unit was then placed in a float rack in a beaker containing 1000 ml of H<sub>2</sub>O and allowed to dialyze with gentle mixing for 1 hour at room temperature. The dialyzed, ligated DNA was then electroporated into *E. coli* DH10B cells (Hanahan, Jessee et al. 1991) using a combination unit made of the Pulse Controller Plus ® and Gene Pulse II ® (BioRad,

Hercules, CA). Five  $\mu$ l of ligated DNA were added to 40  $\mu$ l of *E. coli* DH10B cells and pulsed at 2.5 kV and 25  $\mu$ FD at 200 Ohms. Electroporated cells were then transferred to 960  $\mu$ l of SOC medium for recovery for 45 minutes at 37<sup>0</sup> C. Cells were then plated onto L agar supplemented with kanamycin, x-gal and chloramphenicol as previously described. Ten random white colonies were selected for further analysis and were grown overnight with shaking at 37<sup>0</sup> C in terrific broth with kanamycin and chloramphenicol as described earlier.

The plasmid DNA was then purified from the different clones using the Quantum Prep Plasmid Miniprep ® kit (BioRad, Hercules, CA) following the manufacturer's instructions. Plasmid DNA was then double digested using *Eco*R1 and *Hind*III (both from Invitrogen, Carlsbad, CA) as directed by manufacturer. Digested DNA was separated by electrophoresis in 1x Tris borate-EDTA electrophoresis buffer (TBE) (Sambrook and Russell 2001) on a 0.7% agarose gel (Invitrogen, Carlsbad, CA) and stained using ethidium bromide (5  $\mu$ g/ml). Based on gel results, selected clones were chosen for sequencing. These clones were again grown in terrific broth overnight and their plasmids were isolated using the QIAfilter ® plasmid maxi prep kit (Qiagen, Valencia, CA). The maxi prep kit was chosen instead of the recommended midi prep (recommendation based on volume) because the cells had been grown in terrific broth, which was not recommended by the manufacturer. The manufacturer's recommendations were used with the following exceptions. Cells were harvested at 8,000 g for 5 minutes not the suggested 6,000 g for 15 minutes. Also, the speedvac system was used to dry pellets instead of the recommended air-drying. Pellets were resuspended in 400  $\mu$ l of TE.

The amount of DNA in each clone sample was determined by electrophoresis alongside of a High DNA Mass ® Ladder (Invitrogen, Carlsbad, CA). The remaining DNA was ethanol precipitated, resuspended at a concentration of 200 ng/µl (Maniatis, Fritsch et al. 1982) and sent for sequencing (Virginia Commonwealth University, Massey Cancer Center, Nucleic Acids Research Facilities, Richmond, VA) using the primers listed in Table 1.

Table 1. Primers used in this study

Primer Name	Experiment(s) Used In	Sequence
M13F	1	CGCCAGGGTTTCCCAGTCACGAC
M13R	1	AACAGCTATGACCATG
M29L19	1	AGCGACGCCATCTATGTGT
M1201U19	1,2	TCGGGTATCGCTCTTGAAG
Arb 1-1	2	GGCCACGCGTCGACTAGTCA(N10) AGCTG
Arb 1-2	2	GGCCACGCGTCGACTAGTCA(N10)TGAAC
Arb 2	2	GGCCACGCGTCGACTAGTCA
M110L14	2	AGCCCGGGAATCAT
425L19	2	TTTTCGTTTGTGAACCAT
M1186U16	2	TGAGATAATGCCGACT
M1191U18	2	TAATGCCGACTGTACTTT
M51L19	2	CTAGCGACGCCATCTATGT
IR In	1	AGGTTGGCTGATAAGTC
IR Out	2	GACTTATCAGCCAACCT
1544L19	2, 3	TGTCATGCCATCCGTAAGA
1551L19	2, 3	CTCTTACTGTCATGCCATC
p1894L16	2, 3	ACAGGTTTCCCGACTG
p2075L16	2, 3	TTTCCTGCGTTATCCC
p2280L16	2, 3	CGGCAGGGTCGGAACA
p2532L16	2, 3	ACCGCTACATACCTC
p3013L16	2, 3	CGGCCTATTGGTTAAA
p3507L19	2, 3	GGCGTTACCCAACCTTAATC
p3574L19	2, 3	CGCCCTATAGTGAGTCGTA
p1554U17	2, 3	CTGCGGCCAACTTACTT
p1733U18	2, 3	ATAGCTTGGCGTAATCAT
p2021U15	2, 3	GGCGAGCGGTATCAG
p2047U18	2, 3	GCGGTAATACGGTTATCC
p2446U19	2, 3	TCCGGTAACATCGTCTTG
p3330U17	2, 3	GGCGCTGGCAAGTGTAG
p3495U15	2, 3	TGCTGCAAGGCGATT
p3503U19	2, 3	GGCGATTAAGTTGGGTAAAC
1544L19	2, 3	TGTCATGCCATCCGTAAGA
PGP258U16	3	AGCTGGTCACGACGAA
PGP414U16	3	GCTTAGGCGGATTGTG
PGP433U16	3	AAGGGTGGTCGGAAC
PGP2678U19	3	AGGCCCTCAAGGTGTCTAC

Table 1. Continued

Primer Name	Experiment(s) Used In	Sequence
PGP2578U17	3	TTAGGTGGAACACTCTT
PGP4410L16	3	CCATTCGCCGATGTCA
PGP4698L16	3	CCCGGACGCATACGAA
PGP4966L16	3	ACCTCCCGATACCTTG
PGP2812L16	3	CCACAATCCGCCTAAG
PGP2769L17	3	CAGGGTGACTCAACTCT
PBP6092L16	3	TCCGGCCAGTCGTAAA
PBP6296L18	3	GATAATATCGCGAGTGAT
PBP8091L15	3	CTGCGTCCGCCATAC
PBP8259L19	3	AACTGTGGCATTGGTCATC
PBP8141L15	3	CAGCCCGCATGACGC
PBP6894U19	3	ACGACTGGCCGGACTATTT
PBP7474U14	3	TCACGCGATCAATC
Dala5430U16	3	AAAATCGGCGGAATTA
Dala5715U16	3	GCGGGAAGTTTCGGTT
Dala5899U16	3	AAGCCCAGCGATATTT
Sc447-807U16	3	CAAGGCAAGCGAGCTG
Sc447-981U18	3	GGCAATGCGACCTACTCT
Sc270-406U19	3	GGGCAGAGGACTTGATTGT
Sc270-446U18	3	AGTTCGTGATGCGGAGAC
Sc270-465U19	3	GGCAAGGTGCTGACGGACT
GTP2134L18	3	AGATTGGGCACGATTGTC
GTP2489L16	3	GGAAGATTTGCGGTAG
GMP89U16	3	GGCTGCCGTAATCAAG
GMP389U16	3	ATGTTACCCGCATATA
PGM19071U16	3	GGTTTCACGACGACTT
PGM19204U16	3	CAGTAGGGCATATCCA
PGM23333L14	3	TACGCCCTATTAAC
PGM23478L15	3	CGGCTCCTCTACACT

1 = Clone Sequence Analysis

2 = AP-PCR

3 = Specific PCR Experiments

### Mutant Characterization by AP-PCR

Arbitrarily Primed Polymerase Chain Reaction (AP-PCR) uses degenerate primers paired with transposon-specific primers in order to characterize the DNA flanking the inserted mini-transposon. It is a modification of a method provided by Dr. Glen Tamura (University of Washington, Seattle, Washington). Primers were designed using OLIGO ® Primer Analysis Software (Molecular Biology Insights, Inc., Cascade, CO). Primer sequences can be found in Table 1. The PCR portion of the protocol was done in two steps. The first round of PCR analysis used 8 reactions as shown in Table 2. Each of the arbitrary primers (Arb) was used at 100 pmol/µl and all other primers are used at 25 pmol/µl.

Each first round reaction consisted of 44 µl of Platinum ® PCR Supermix (Invitrogen, Carlsbad, CA), 2 µl of mutant chromosomal DNA at 250-500 ng/µl, 1 µl of primer 1, 1 µl of primer 2 and 2 µl of H<sub>2</sub>O. The reaction was run under the following conditions: 95°C for 5 minutes followed by 6 cycles (95°C x 30 seconds, 30°C x 30 seconds then 72°C for 1.5 minutes) then 30 cycles (95°C x 30 seconds, 45°C x 30 seconds then 72°C for 2 minutes) then 72°C for an additional 4 minutes and finally cooled to 4°C. Reactions were run on a GeneAmp ® PCR System 2400 (Perkin Elmer, Shelton, CN). A 10 µl sample from each of the first round reactions was separated by electrophoresis and if products were detected, the remaining PCR product was purified using QIAquick® PCR purification kit (Qiagen, Valencia, CA). The PCR products were eluted with 30 µl of the provided buffer.



Table 2. Primer Pairs used for AP-PCR

Reaction	Template	Primer 1	Primer 2
<b>First Round Reaction Pairs</b>			
1	Mutant DNA	Arb1-1	M110L14 (L)
2	Mutant DNA	Arb1-1	425L19 (L)
3	Mutant DNA	Arb1-1	M1186U16 (R)
4	Mutant DNA	Arb1-1	M1191U18 (R)
5	Mutant DNA	Arb1-2	M110L14 (L)
6	Mutant DNA	Arb1-2	425L19 (L)
7	Mutant DNA	Arb1-2	M1186U16 (R)
8	Mutant DNA	Arb1-2	M1191U18 (R)
<b>Second Round Reaction Pairs</b>			
9	1	Arb2	M51L19 (L)
10	2	Arb2	M51L19 (L)
11	3	Arb2	M1201U19(R)
12	4	Arb2	M1201U19(R)
13	5	Arb2	M51L19 (L)
14	6	Arb2	M51L19 (L)
15	7	Arb2	M1201U19(R)
16	8	Arb2	M1201U19(R)

The purified products of the first round reactions were then used as template DNA for the second round reactions. Primer pairs for the second round reactions can also be found in Table 2.

Each second round reaction consisted of 88  $\mu$ l of Platinum ® PCR Supermix (Invitrogen, Carlsbad, CA), 4  $\mu$ l of template DNA (from the corresponding first round reaction), 2  $\mu$ l of primer 1, 2  $\mu$ l of primer 2 and 4  $\mu$ l of H<sub>2</sub>O. The reaction was run under the following conditions: 95°C for 1 minute followed by 30 cycles (95°C x 30 seconds, 52°C x 30 seconds then 72°C for 2 minutes) then 72°C for an additional 4 minutes and finally cooled to 4°C. A 5  $\mu$ l aliquot of the second round product was separated by electrophoresis on a 1.4% agarose gel and stained with 1  $\mu$ l Gelstar (Cambrex, Rockland, ME) per 10 ml buffer. Those samples that had detectable products were then ethanol precipitated (Maniatis, Fritsch et al. 1982). The DNA pellet was reconstituted in 15  $\mu$ l of TE and the entire amount was then separated by electrophoresis on a 1.4% gel stained with Gelstar. If the band of interest was very bright on the initial electrophoresis, the sample was diluted 1:1 with H<sub>2</sub>O and run out in two lanes during the second electrophoresis. A gel purification procedure was then done on all bands of interest using MinElute® Gel Extraction kit (Qiagen, Valencia, CA) following manufacturer's directions. Recovered DNA was eluted with 16  $\mu$ l of H<sub>2</sub>O.

The DNA concentration of each sample was determined by running a 2  $\mu$ l aliquot on a gel alongside 2  $\mu$ l of Low DNA Mass ® Ladder (Invitrogen, Carlsbad, CA). The remaining sample was then adjusted to a concentration of 10 ng/ $\mu$ l and sent for

sequencing using the same primers used in the second round reaction above. Primer concentrations for sequencing were 3.2 pmol/ $\mu$ l.

### Polymerase Chain Reaction

Numerous polymerase chain reaction (PCR) analyses were done using a variety of primers to determine the sequence of DNA flanking the inserted transposon. The primers used are listed in Table 1. The annealing temperature was adjusted to be approximately 5° below the  $T_m$  of the primer with the lowest  $T_m$ . The extension time was generally 2 minutes; however, it was lengthened to 1 minute per kb of expected product length when necessary. The general conditions for PCR were 94°C for 4 minutes followed by 30 cycles (94°C x 30 seconds,  $T_m$ -5°C x 30 seconds then 72°C for 2 minutes) then 72°C for an additional 4 minutes and finally cooled to 4°C.

For amplification of signature tags the conditions were 95°C for 1 minute followed by 25 cycles (94°C x 30 seconds, 45°C x 1 minute then 72°C for 10 seconds) then 72°C for an additional 2 minutes and finally cooled to 4°C. Products were separated by electrophoresis in 3% Nusieve ® 3:1 agarose (BioWhittaker Molecular Applications, Rockland, ME).

### DNA Gel Electrophoresis and Southern Blotting

Gel electrophoresis was performed using 0.7% agarose if the expected band size was greater than 1 kb or 1.4 % agarose (Invitrogen, Carlsbad, CA) if the expected band size was between 100 bp and 1 kb. For bands expected to be smaller than 100 bp, 3% Nusieve ® 3:1 agarose (BioWhittaker Molecular Applications, Rockland, ME) was used. All gels were run using the Sub-cell <sup>TM</sup> GT Electrophoresis Cell system with the

PowerPac<sup>TM</sup> 300 power supply system (BioRad, Hercules CA). DNA was visualized with either 1 µl Gelstar per 10 ml of agarose (Cambrex, Rockland, ME) or with ethidium bromide (0.2 µg/ml).

Southern blot experiments were done using the Genius digoxigenin system (Roche Molecular Biochemicals, Indianapolis, IN). Digoxigenin-d-UTP was incorporated into probes by PCR labeling and detected with anti-digoxigenin antibodies conjugated to alkaline phosphatase. Chemiluminescence was produced by addition of the substrate CDP Star according to the manufacturer's instructions (Roche Molecular Biochemicals). Images were recorded using a FluorChem 8000 digital imaging system (Alpha Innotech, Inc., San Leandro, CA). The following restriction enzymes were used: *EcoRI* (Invitrogen, Carlsbad, CA), *ScaI*, *AccI* and *BspEI* (New England Biolabs, Inc., Beverly, MA). Southern blot results were analyzed with the aid of Gene Construction Kit<sup>TM</sup> 2.5 (Textco, Inc., Research Triangle Park, NC).

#### DNA Sequence Analysis

Sequencing was performed at Virginia Commonwealth University, Massey Cancer Center, Nucleic Acids Research Facilities, Richmond, VA. DNA sequences were assembled using SeqMan<sup>TM</sup> II (DNASTAR, Inc., Madison, WI). Gene homology was determined by searching GenBank databases using BLASTN and BLASTX programs (Altschul, Madden et al. 1997).

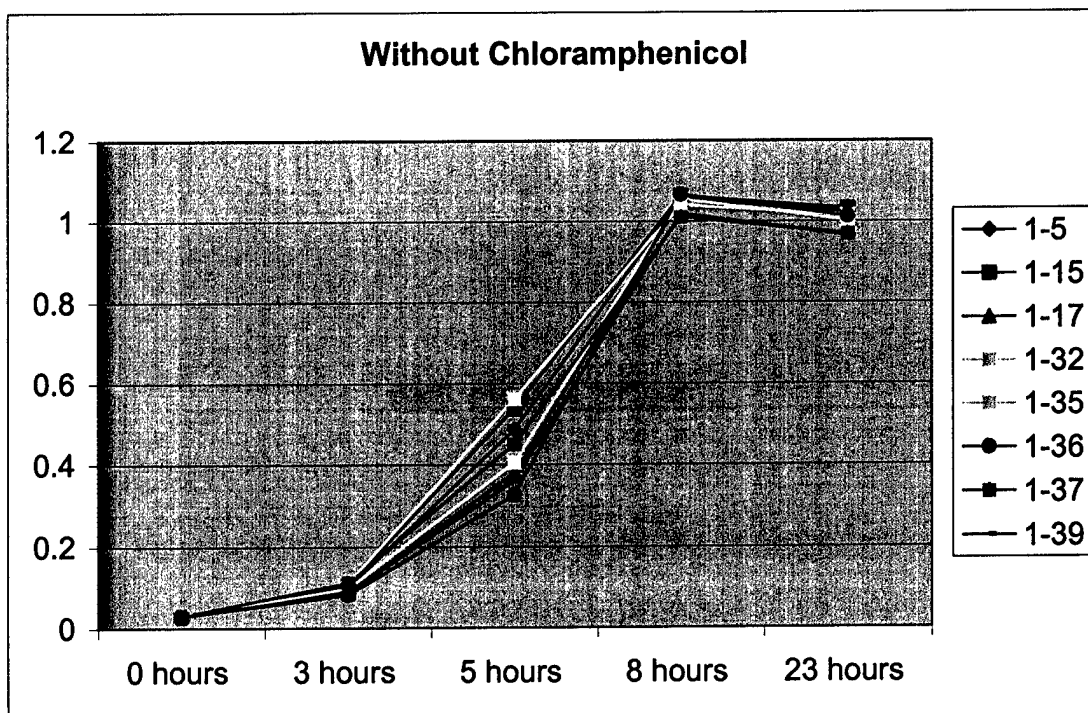
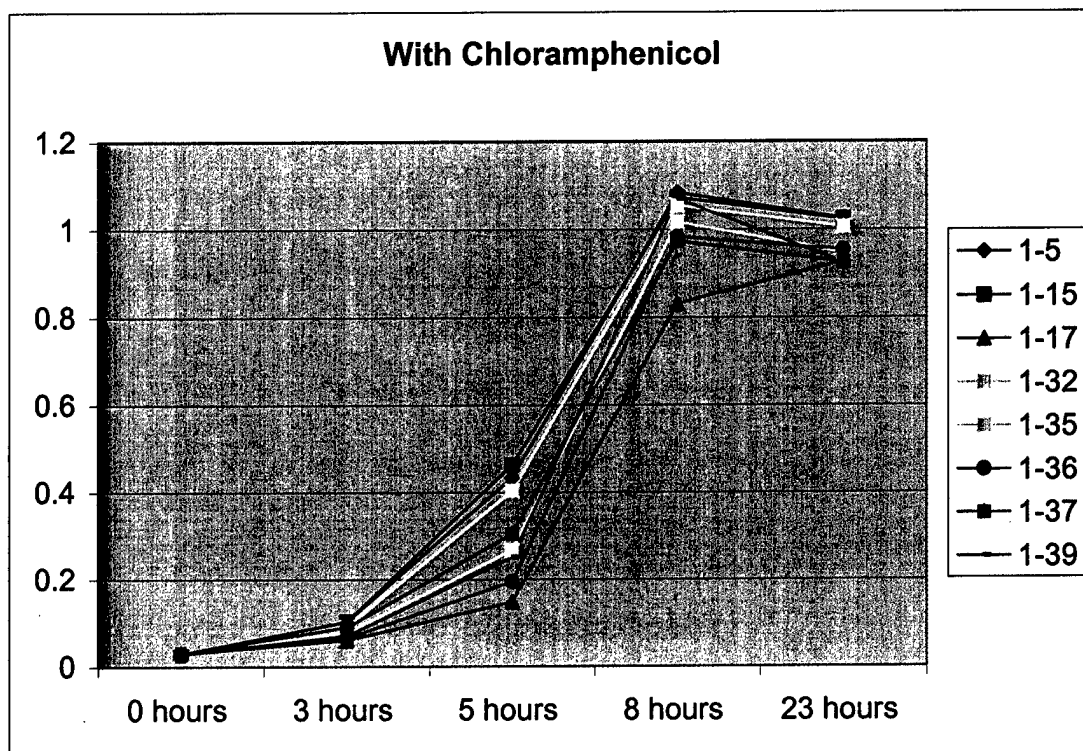
## Results

### Growth Studies

Growth studies were performed to determine whether poor growth of the mutant strains could be the cause of the decreased signal seen when comparing the results of the input pool to the output pool (Fig. 4). To determine this, overnight cultures were grown in 5 ml of BHI with 5 µg/ml of chloramphenicol. The overnight cultures were used to inoculate fresh media. Densities of the overnight cultures were determined by optical density at 660nm (OD<sub>660</sub>) and the inocula matched by dilution with additional BHI. Cultures were incubated in both the presence and absence of 5 µg/ml chloramphenicol and the OD measured at various times (3, 5, 8 and 23 hours). Results of the growth studies can be seen in Figure 4. The mutants that appeared avirulent exhibited somewhat reduced growth compared to the normal controls at 5 hrs, though growth in all strains was essentially the same by 8-24 hrs. Thus it seemed possible that poor growth may have contributed to the decreased signal detected in the STM experiment.

Figure 4. Growth Studies.

Measurement of the OD<sub>660</sub> at 0, 3, 5, 8 and 23 hours was measured. Strains in red (1-17, 1-36, 1-37 and 1-39) had a decreased signal in the output pool. Strains in yellow (1-32 and 1-35) had a decrease in signal on the inoculum blot. Strains in blue (1-5 and 1-15) served as normal controls.



### DNA Isolation and Quantification

DNA from mutant strains 1-5, 1-17, 1-32, 1-35, 1-36, 1-37 and 1-39 was isolated and quantified as described previously. The DNA was isolated and the concentration determined for use in experiments and to ensure that similar DNA yields could be found in all strains. Results of the quantification procedure can be seen in Table 3. DNA concentration was determined from the following formula:

$$\text{DNA Concentration} = \text{OD}_{260} \times 50 \mu\text{g/ml} \times 50 \text{ (dilution factor)}$$

The ratio of  $\text{OD}_{260}:\text{OD}_{280}$  is an indication of the purity of the DNA preparation. Ideally the ratio should be 1.8. Results that are significantly higher or lower have a larger amount of impurities such as RNA and/or protein.



**Table 3.      DNA Quantification Results.**

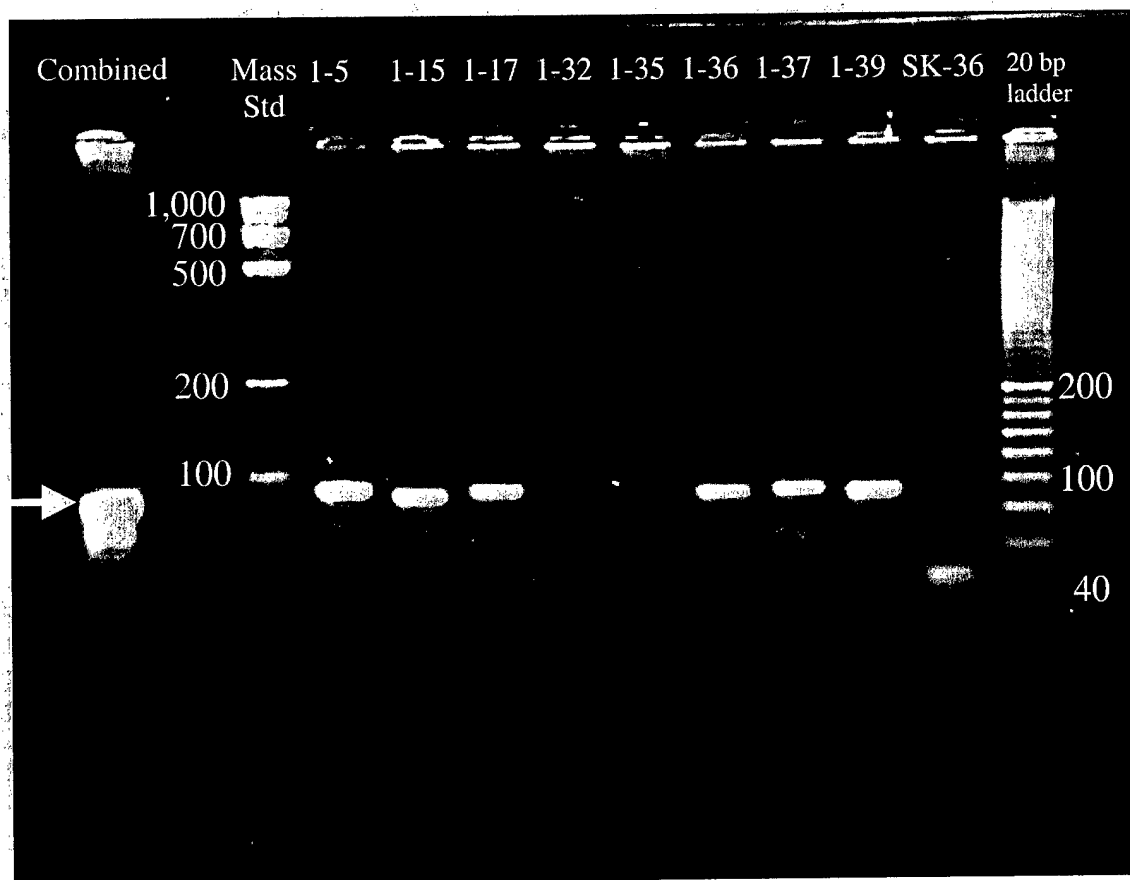
<b>Sample</b>	<b>OD<sub>260nm</sub></b>	<b>OD<sub>280nm</sub></b>	<b>Ratio OD<sub>260</sub>/OD<sub>280</sub></b>	<b>DNA Concentration (µg/ml)</b>
1-5	0.211	0.115	1.834	527.5
1-15	0.203	0.108	1.879	507.5
1-17	0.199	0.121	1.644	497.5
1-32	0.285	0.161	1.770	712.5
1-35	0.205	0.108	1.898	512.5
1-36	0.246	0.135	1.822	615.0
1-37	0.206	0.110	1.872	515.0
1-39	0.198	0.105	1.885	495.0

### Amplification of Signature Tags

Each of the tags from mutant strains 1-5, 1-15, 1-17, 1-32, 1-35, 1-36, 1-37, 1-39 and parent strain SK36 were PCR amplified, ethanol precipitated and separated by electrophoresis. The signature-tag amplification was done to determine if the cause of the decreased signal strength was due to a problem in the signature tag itself. Results showed that most tags amplified as expected with the exception of 1-32 and 1-35 (Figure 5). These are the same mutants that failed to show adequate signal strength on the inoculum dot blot. It therefore appeared that failure of signature tag amplification could explain the weak signals shown by mutants 1-32 and 1-35 in all of the blots, but not the decreased signal exhibited by mutants 1-17, 1-36, 1-37, and 1-39 in the output blots (Figure 3).

Figure 5.      Signature-Tag Amplification

Signature tags were amplified by PCR from the strains indicated, separated by electrophoresis on a 3 % agarose gel, and visualized by ethidium bromide staining. The combined tube "Comb" shows the PCR products obtained from an amplification reaction in which equal amounts of genomic DNA from all eight mutant strains was used as a template. SK36, the parent strain for the mutants was used as a negative control.



### Southern Blot Results

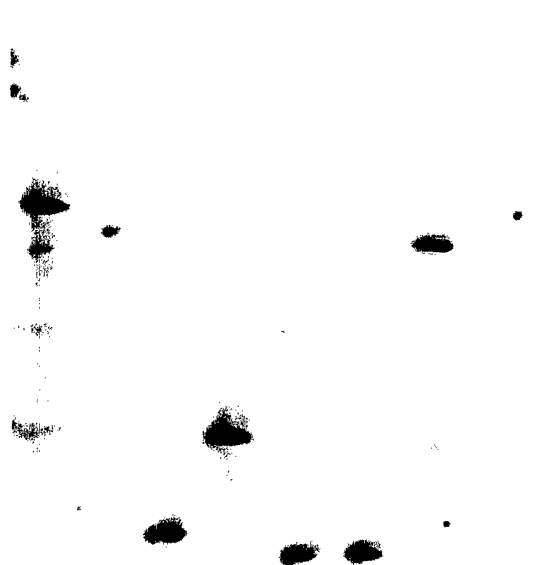
A Southern blot was performed using *Eco*RI-digested chromosomal DNA from mutant strains 1-5, 1-15, 1-32, 1-35, 1-36, 1-37, 1-39 and the parent strain SK36. Probes were made from *Eco*R1 digested pJFP1 (which includes the mini-transposon) and from the PCR amplified transposon. When using either probe, one to two bands would be expected depending on the distance between the nearest *Eco*R1 site in the flanking chromosomal DNA and the site within the transposon of each mutant. Unexpectedly, 3 bands were observed in some of the mutants (1-32, 1-36, 1-37 and 1-39) when probed with the whole plasmid. However, only one or two bands were seen in the mutants when probed with the amplified mini-transposon.

Figure 6.      Southern Blots of Mutant Strains

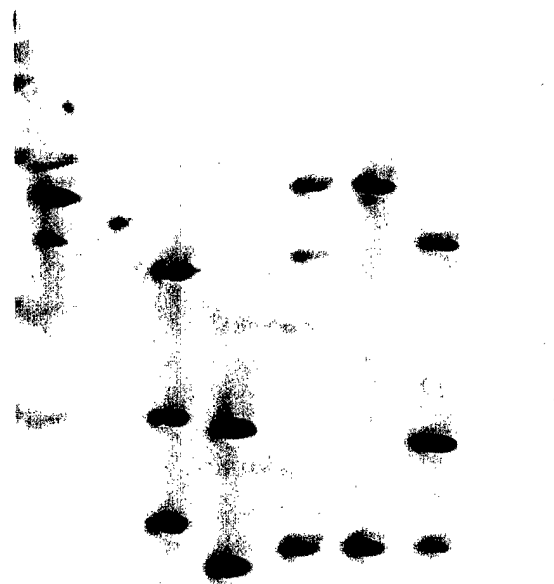
Chromosomal DNA from the strains indicated was digested with *Eco*RI, separated by electrophoresis on a 0.7 % gel, and transferred to a nylon membrane. Following hybridization and stringent washing, bound probe DNA was detected by chemiluminescence. Panel A: PCR amplified mini-transposon DNA used as the probe. Panel B: *Eco*RI digested pJFP1 plasmid DNA used as the probe.

A.

1-5 1-17 1-32 1-35 1-36 1-37 1-39 SK36

B.

1-5 1-17 1-32 1-35 1-36 1-37 1-39 SK36



### Characterization of Mutant 1-39

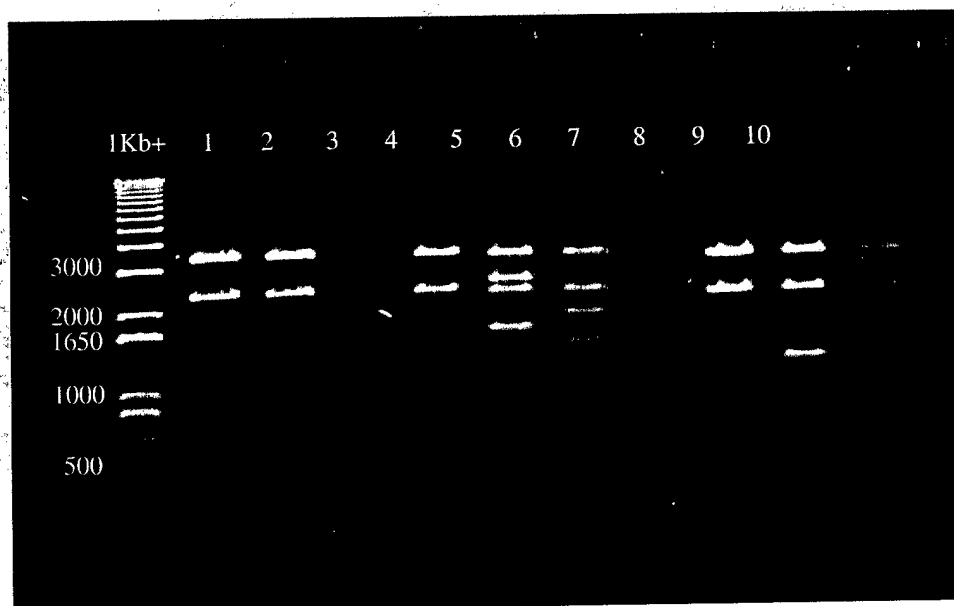
In order to identify the putative virulence gene disrupted by the insertion of the signature tag in mutant 1-39, we first cloned the transposon with flanking DNA from the mutant into a plasmid, selecting for chloramphenicol resistance in *E. coli*. Ten random white colonies were picked and their DNA doubly digested with both *Eco*R1 and *Hind*III. The results are shown in Figure 7A. Clones 3, 7 and 10 had insufficient DNA for further testing. All clones except #6 appeared to possess multiple common bands. This would be the expected result if the various clones all contained the transposon in addition to varying lengths of flanking DNA. Next, clones 1, 2, 4, 5, 6, 8 and 9 were digested with only *Eco*RI (Figure 7B). Clones 1 and 8 appeared to be identical as did clones 2 and 4. All clones except number three had the expected 3.3 kb vector band visible. Based on these results it was determined that clones 1 and 2 would be grown overnight for plasmid isolation. The amount of DNA in each clone sample was determined by electrophoresis in a 0.6% agarose gel visualized with ethidium bromide alongside of a High DNA Mass ® Ladder. The remaining DNA was ethanol precipitated and sequenced.



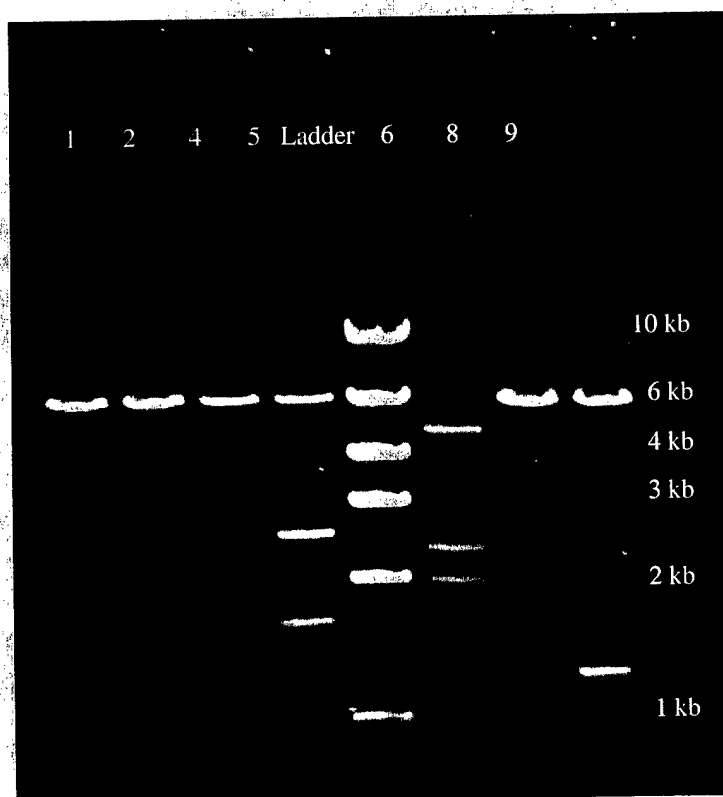
Figure 7. 1-39 Cloning Experiment.

Mutant 1-39 DNA was partially digested using *Sau3AI* then ligated to vector plasmid pVA2606 that had been previously digested with *BamHI*. The ligated DNA construct was electroporated into *E. coli* DH10B cells and grown on L agar supplemented with kanamycin, x-gal and chloramphenicol to select for the desired transformants. Ten randomly picked white colonies were selected for DNA isolation. Recombinant plasmids potentially containing the transposon and flanking DNA were either doubly digested with *EcoRI* and *HindIII* (A) or digested only with *EcoRI* (B). The DNA from each was then separated by electrophoresis on a 0.7 % agarose gel, and visualized by ethidium bromide staining.

A.



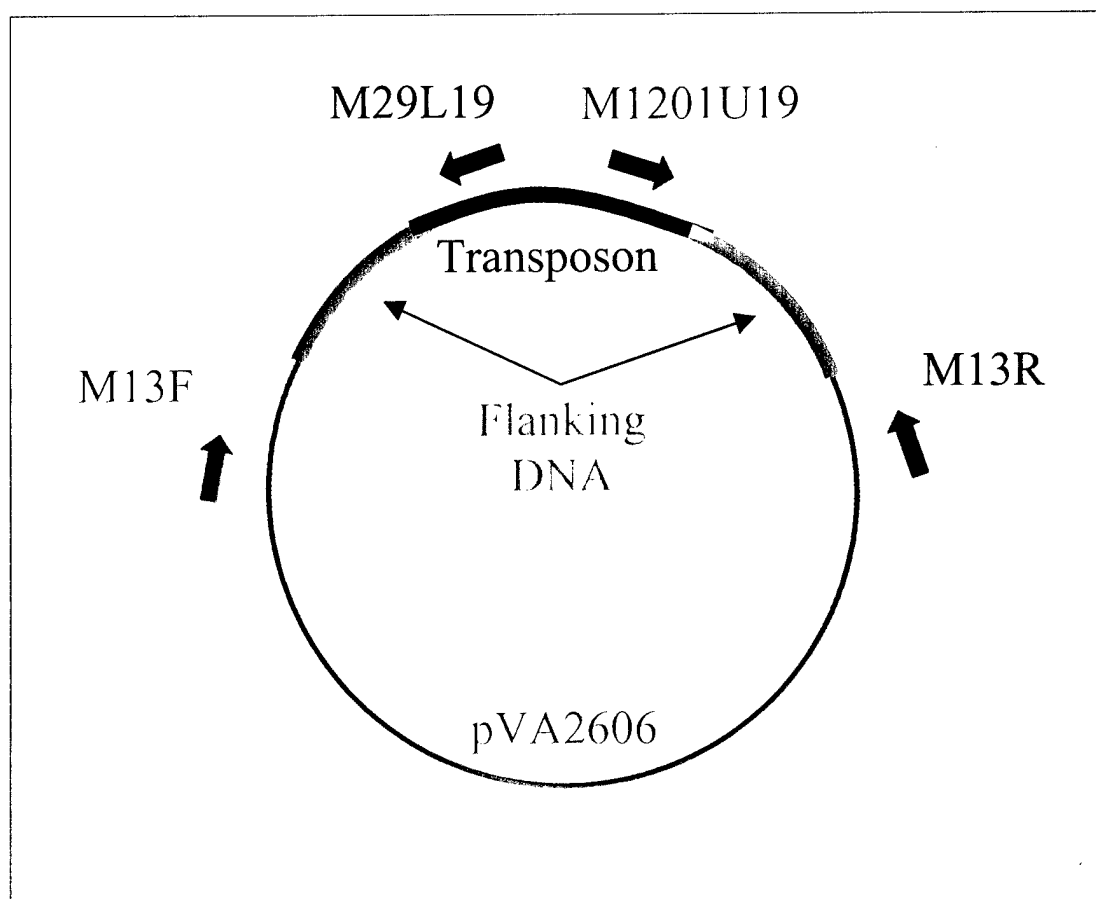
B.



Sequence results were assembled using the Seqman II program and gene homology determined by searching GenBank databases. The relative position of the sequencing primers in relation to the transposon, the flanking DNA and the vector DNA is shown in Figure 8. Sequence data obtained showed that DNA flanking the right of the transposon had homology to the penicillin binding protein 2B (*pbp-2B*) gene from *Streptococcus pneumoniae*. Surprisingly, left-flanking sequences had homology to the plasmid pJFP1 initially used in the *in vitro* transposition reaction step of STM. Further analysis showed that the plasmid pJFP1 has a multitude of *Sau3AI* restriction digest cut sites and since it appears to have integrated into the mutant genome along with the mini-transposon, we decided to abandon cloning as the method to characterize the left flanking sequence of the inserted DNA (pJFP1 + mini-transposon).

Figure 8. Position of Sequencing Primers used in Cloning, 1-39

Four different primers were used to sequence mutant 1-39 clones #1 and #2. The forward and reverse primers M13F and M13R from the cloning vector pVA2606 were used as well as two others the transposon. M29L19 anneals to the 5' end of the transposon and faces to the left and M1201U19 anneals to the 3' end and faces to the right.



Next, AP-PCR was used to characterize the DNA flanking the insertion. AP-PCR is a relatively new adaptation of the PCR technique. This technique was done in two rounds (Figure 9). In the first round, a transposon-specific primer (or plasmid-specific primer if the plasmid pJFP1 inserted with the transposon) was paired with one of two degenerate primers (Arb 1-1 or Arb 1-2). Primer sequences can be found in Table 1. The two first round arbitrary primers have different 3' sequences and a common 5' sequence. If a product was detected in the first round reaction it was used as a template for the second round reaction. The second round reaction used a transposon (or plasmid) primer that was nested in relation to the first round transposon primer and paired with the second round arbitrary primer (Arb 2), which is composed of only the 5' common portion of Arb 1-1 and Arb 1-2. For the first round PCR reactions, six cycles were run with an annealing temperature of 30° C to allow for relaxed annealing and then raised to 45° C for an additional 30 cycles for more specific binding. Since the second round reactions use specific primers, the annealing temperature is raised even further to 52° C for 30 cycles for greater specificity.

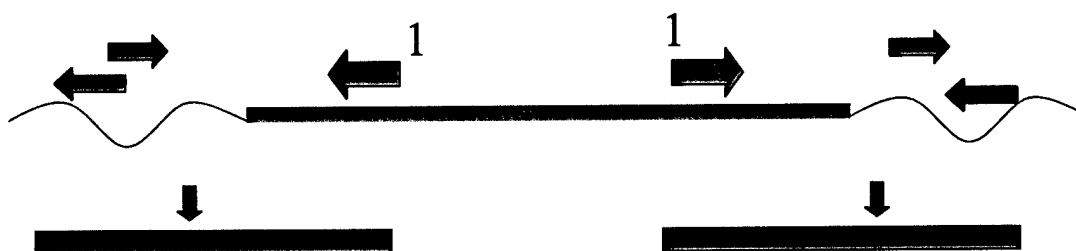
Figure 9. Principle of AP-PCR

In this technique, primers containing random sequence near their 3' ends (Arb1-1) are paired with a primer (labeled "1") that binds within the transposon. A first round reaction is done using these primers. If product is detected, a second round reaction is performed using primer Arb-2, which corresponds to the constant 5' portion of Arb1-1, along with a nested primer (labeled "2")- that is, a primer that is "inside" the desired product from the first PCR. Products detected from the second round reaction are sequenced using the primers from the second round reactions.

- First performed PCR using random primers paired with transposon specific primers:

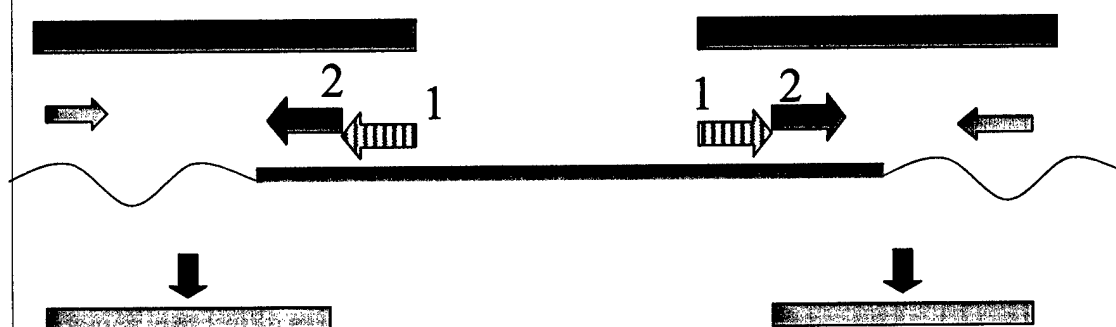
–Arb1-1

GGCCACGCGTCGACTAGTCA(N10)AGCTG



- Then if product was detected – performed second round PCR using nested primers and:

–Arb 2 GGCCACGCGTCGACTAGTCA





After numerous attempts with various primer pairs at several different concentrations, we optimized the protocol and confirmed that plasmid DNA was indeed inserted along with the transposon. We also confirmed that the *pbp*-2B gene was located to the right of the insertion. Additionally, we determined that the DNA flanking the left of the plasmid had homology to the phosphoglycolate phosphatase gene (*pgp*) of *S. pneumoniae*.

Based on these findings, primers were designed from both the flanking *pbp* and *pgp* sequences. These primers were then used in conjunction with a variety of primers from within the transposon and the plasmid to confirm our findings. We successfully amplified products using both *pgp* and *pbp* primers matched to primers from within the plasmid and the mini-transposon. These results indicated that a partial copy of the *pgp* gene was located to the left (5') of the vector and transposon insertion and that a partial copy of the *pbp*-2b gene was located to the right (3') as expected. Figure 10 shows an overview of the 1-39 sequences that were assembled with the program Seqman II. Next, we searched against the ongoing *S. sanguis* sequencing project contigs and found matches to both the right and left flanking sequences. A portion of contig 342 containing the *pbp*-2b gene is shown in Figure 10.

Because the *pbp*-2b and *pgp* sequences were found within separate, large contigs, it appeared unlikely that these two genes were adjacent to one another in the SK36 genome. Since a simple insertion would have separated chromosomal sequences that were previously contiguous, this finding raised questions as to the nature of the insertion event that occurred in mutant 1-39. To investigate this question further, primers

were designed from the matching contigs to determine whether intact copies of either gene could be amplified.

Surprisingly, both genes were successfully amplified, suggesting that intact copies of both genes were present in addition to the partial copies found adjacent to the insertion. This finding raised more questions than answers. To investigate further, Southern blotting was performed using three separate probes made from the *pgp* gene, the transposon, and the right flanking portion of the *pbp-2b* gene. These probes were hybridized with three identical blots, containing SK36 (parental) and 1-39 mutant DNA digested with each of three restriction enzymes, (*EcoRI*, *ScaI* and *BspEI*). The results of the Southern blots (Figure 11) indicate that both a wild type copy and a disrupted copy of each gene (*pgp* and *pbp-2b*) are present in the mutant. It also confirms that only one copy of each gene is present in the parent strain SK36.

Figure 10. Strategy View of Mutant 1-39

DNA sequences aligned with the program Seqman II are displayed graphically. Colors indicate the origin of each sequence. Sequences derived from cloning experiments (green), specific PCR experiments using primers designed from newly determined flanking DNA sequences (pink) and AP-PCR experiments (not highlighted) were assembled along with the mini-transposon (blue) and the vector plasmid pJFP1 (yellow). Additionally, alignment is shown with a contig sequence from the ongoing *S. sanguis* sequencing project containing the *pbp-2b* gene (contig 342; gray).

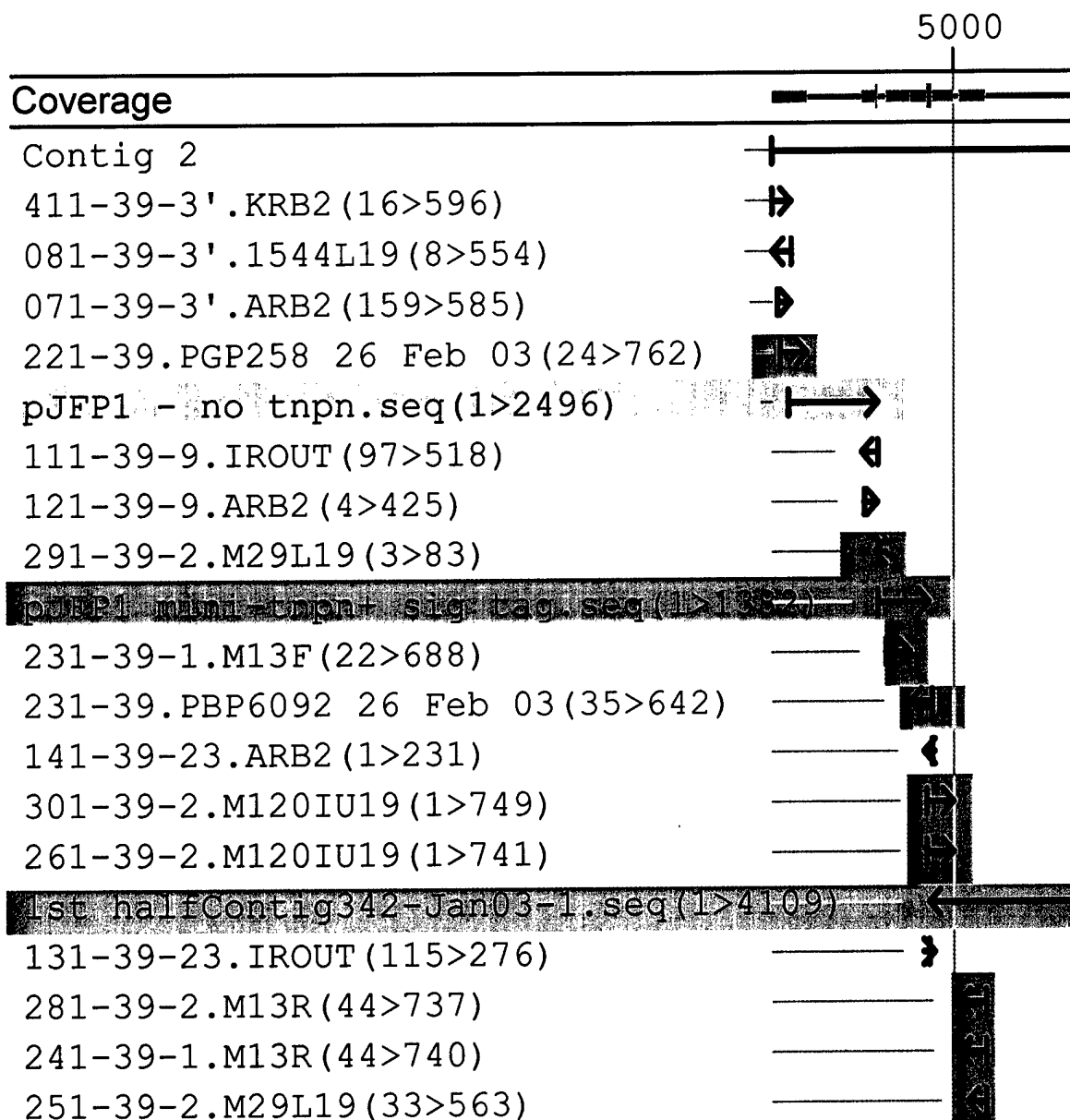
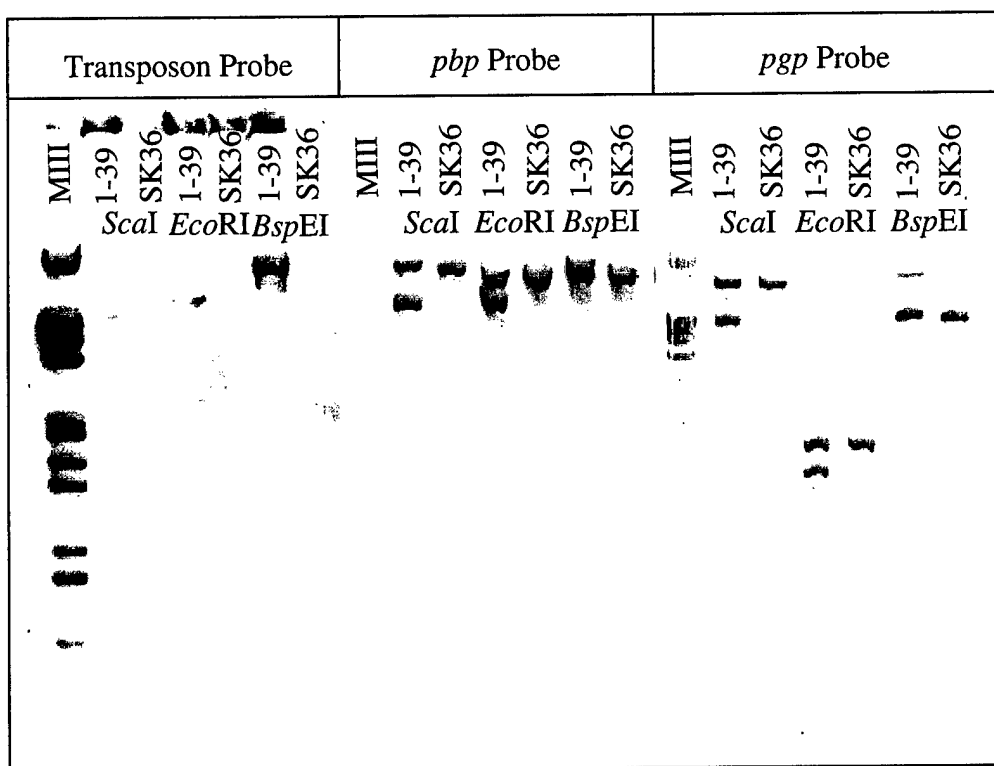


Figure 11. Mutant 1-39, Southern Blot

Chromosomal DNA from the parent strain SK36 and the 1-39 mutant were digested with each of three enzymes, *EcoRI*, *ScaI* and *BspEI*, separated by electrophoresis on a 0.6 % gel, and transferred to a nylon membrane. Following hybridization and stringent washing, 3 different sets of bound probe DNA were detected by chemiluminescence. The digestion sites for *EcoRI* and *ScaI* with-in pJFP1 with the mini-transposon can be seen in Figure 1. There is no digestion site for *BspEI* with in the plasmid.



### Characterization of Mutant 1-17

Next we attempted to identify the putative virulence gene disrupted by insertion of the signature-tagged transposon into mutant 1-17. Analysis of the initial Southern blots probed with the plasmid pJFP1 containing the mini-transposon and with the mini-transposon alone (Figure 6) indicated that only the transposon was inserted into the genome. That is, since there was only one band present in both blots, it appeared that the plasmid had not inserted with the vector in this mutant. Cloning of the transposon and flanking DNA from the mutant into a plasmid and then selecting for chloramphenicol resistance in *E. coli* was attempted without success. Since the AP-PCR procedure had worked well with mutant 1-39, it was used for the remaining mutants.

Again numerous attempts with various primer pairs at several different concentrations were needed to achieve optimal results. Sequencing of second round reaction amplified DNA confirmed that only transposon DNA was inserted. It also indicated that non-contiguous sequences were again present on either side of the transposon. In the case of 1-17, a gene with homology to a guanosine tri-phosphate binding protein gene (GTP-bp) from *S. pneumoniae* was present on the right. On the left, we initially found a sequence identical to those found to the left of 1-36 and 1-37. This served to point out a problem in the initial choice of primers. The first round transposon-specific left-end primer (M29L19) appeared to be less specific than required and the second round nested transposon-specific primer was from the inverted repeat region and very near the terminus of the transposon. This did not allow us to recover enough

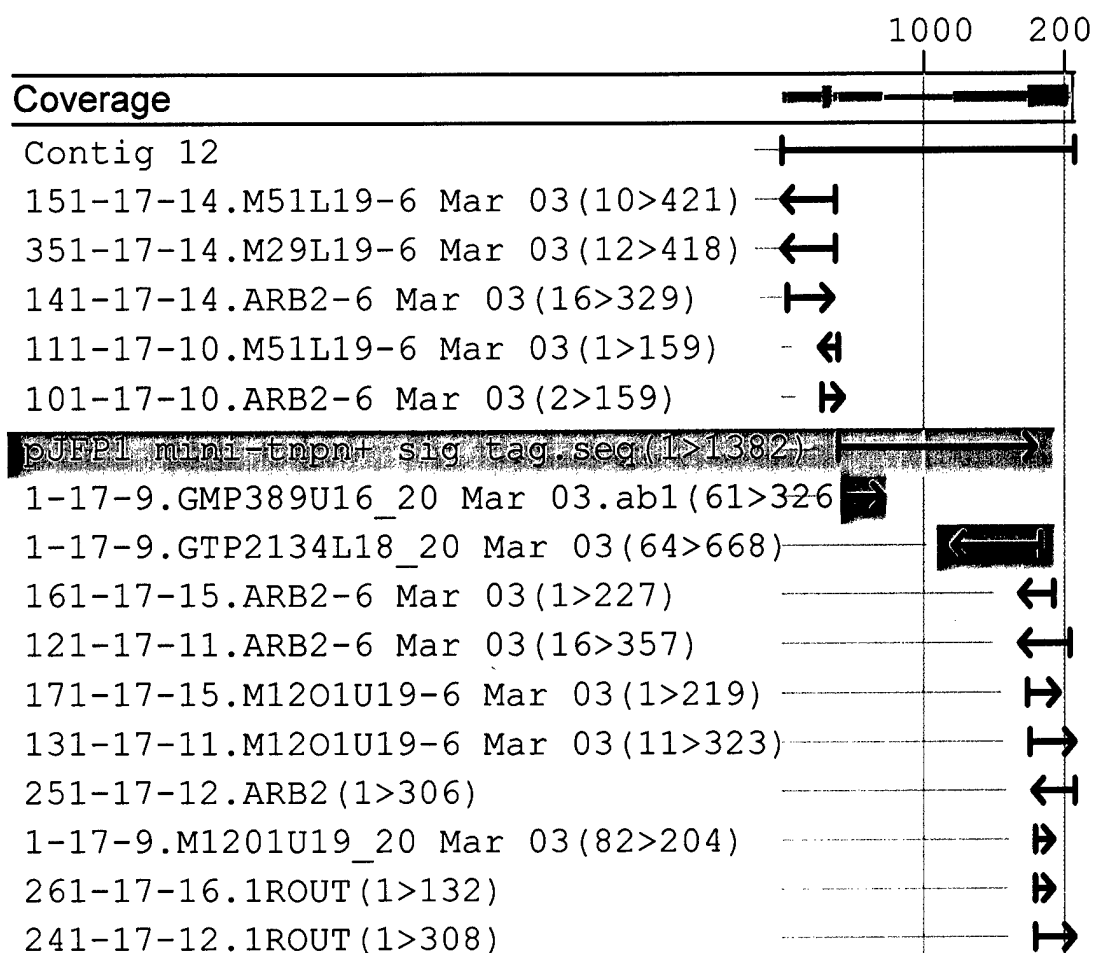
transposon sequence to confirm that the sequence amplified was indeed flanking the transposon. To correct this we designed new first and second round primers for the left end from an area inside of the inverted repeat region of the transposon. Sequence analysis showed that the product of these primers had homology to the guanosine monophosphate synthase gene (GMP-synthase) from *S. pneumoniae*.

Based on these findings, primers were designed from both the flanking GTP-bp and GMP-synthase sequences. These primers were then used in conjunction with a variety of primers from within the transposon to confirm our findings. We successfully amplified products using both GTP-bp and GMP-synthase primers matched to primers from within the transposon. These results indicated that a partial copy of the GMP-synthase gene was located to the left of the transposon insertion and that a partial copy of the GMP-synthase gene was located to the right as expected. Figure 12 shows an overview of the 1-17 sequences that were assembled with the program Seqman II. Next, we searched against the ongoing *S. sanguis* sequencing project contigs and found matches to both the right and left flanking sequences. Although the contig that matched the GMP-synthase gene was approximately 20 kb in length, it was situated such that the end of the contig corresponded to the middle of the disrupted GMP-synthase gene. This contig thus provided no additional information concerning the region upstream from the GMP-synthase gene.



Figure 12     Strategy View of Mutant 1-17

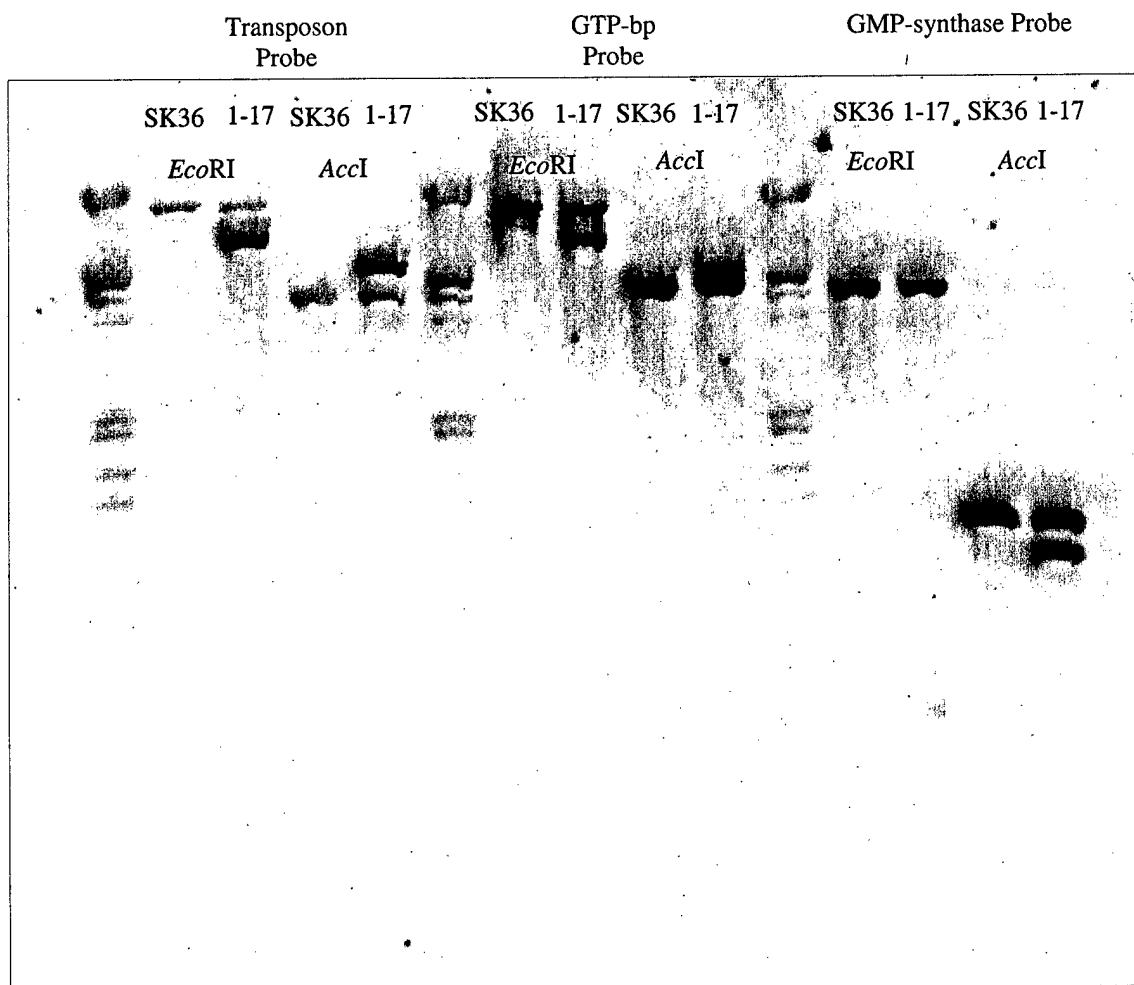
DNA sequences aligned with the program Seqman II are displayed graphically. Colors indicate the origin of each sequence. Sequences derived from specific PCR experiments using primers designed from newly determined flanking DNA sequences (purple) and AP-PCR experiments (not highlighted) were assembled along with the mini-transposon (blue).



As with mutant 1-39, primers were designed from the matching contigs to determine whether intact copies of either gene could be amplified. Again, both genes were successfully amplified, suggesting that intact copies of both genes were present in addition to the partial copies found adjacent to the insertion. Southern blotting was then performed for 1-17 using three separate probes made from the GMP-synthase gene, the transposon, and the GTP-bp gene. These probes were hybridized with three identical blots, each containing SK36 (parental) and 1-39 mutant DNA digested with two restriction enzymes, (*EcoRI* and *AccI*). The results of the Southern blots are shown in Figure 13. The results indicate that both a wild-type copy and a disrupted copy of each gene (GMP-synthase and GTP-bp) were present in the mutant. It also confirmed that only one copy of each gene was present in the parent strain SK36.

Figure 13. Mutant 1-17 Southern Blot

Chromosomal DNA from the parent strain SK36 and the 1-17 mutant were digested with two separate enzymes, *EcoRI* and *AccI*, separated by electrophoresis on a 0.6 % gel, and transferred to a nylon membrane. Following hybridization and stringent washing, 3 different sets of bound probe DNA were detected by chemiluminescence. The appearance of bands in the transposon-probed blot for the parent strain SK36 is probably due to contamination between the different blots during washing.



### Characterization of Mutants 1-36 and 1-37

Next we attempted to identify the putative virulence genes disrupted by the insertion of the transposon into mutants 1-36 and 1-37. Analysis of the initial Southern blots probed with the plasmid pJFP1 containing the mini-transposon and with the mini-transposon alone (Figure 6) indicated that, as in 1-39, both the transposon and the vector plasmid pJFP1 were inserted into the genome in both cases. This was concluded since both mutant chromosomes had three bands when probed with the vector in addition to the transposon. We attempted only AP-PCR on these mutants.

Sequencing of second round AP-PCR amplified DNA confirmed that plasmid DNA was inserted along with transposon DNA. In the case of 1-37, the plasmid was inserted to the right of the transposon (Figure 14A) and in the case of 1-36 there was plasmid DNA on either side of the transposon (Figure 14B). Limited further characterization was done for reasons that will be covered in the discussion; however, in the case of 1-37, left-flanking DNA showed some homology to a putative phosphoglycerate mutase gene (*pgm*) in *S. pneumoniae*. Flanking sequence to the right of the inserted transposon and plasmid was never characterized.

The DNA sequencing results with 1-36 indicated that the transposon was flanked on either side by vector DNA. Attempts to amplify DNA using transposon specific primers paired with various plasmid specific primers showed that only part of the vector was present on each side and that we could not account for the entire plasmid sequence. Furthermore, we were able to determine that DNA flanking the right side, past the vector plasmid sequence, aligned with contig 217 from the *S. sanguis* sequencing project. A

BlastX search was performed using contig 217 and homology was found to iron regulated protein *frpC* from *Neisseria meningitidis*. No further characterization was done for the right side of the insertion.

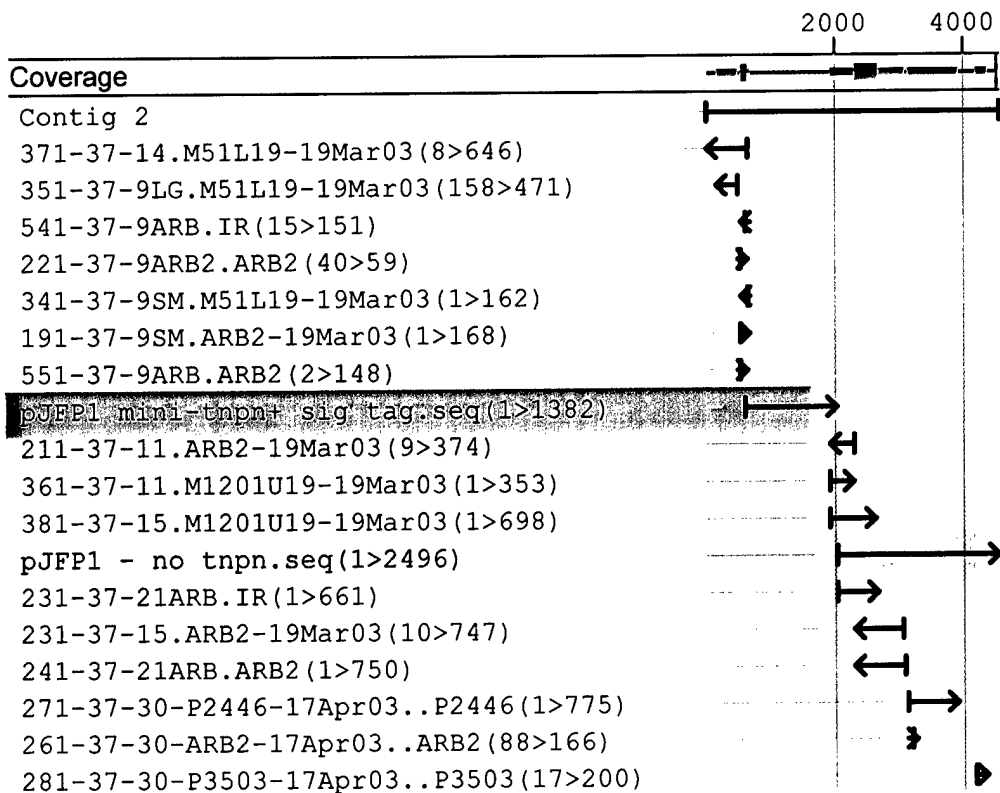
Figure 14     Strategy View of Mutants 1-36 and 1-37

DNA sequences for mutants 1-37 (**A**) and 1-36 (**B**) aligned with the program Seqman II are displayed graphically. Colors indicate the origin of each sequence. Sequences derived from AP-PCR experiments were assembled along with the mini-transposon (blue) and pJFP1 vector plasmid (yellow).

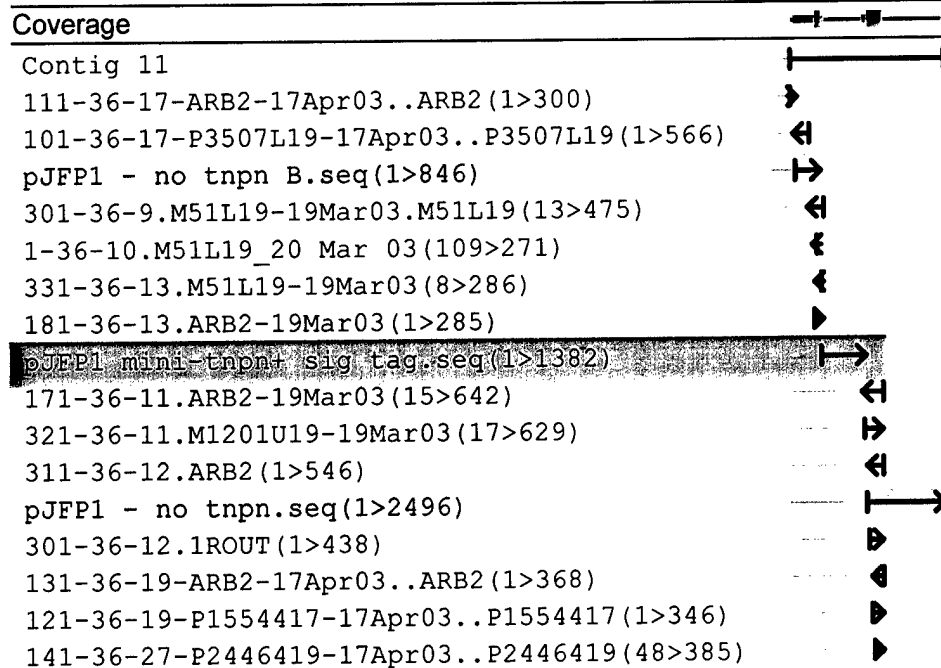


A.

63



B.



## Discussion

The discovery of new virulence genes in *S. sanguis* could lead to the development of a vaccine and to new drug therapies to treat endocarditis. Toward that end, signature-tagged mutagenesis was used to try to identify candidate virulence genes. In this case, *S. sanguis* strain SK36 was mutated such that each mutated strain was "labeled" with a unique signature tag by use of transposon mutagenesis. A mini-transposon from the *mariner* family, in plasmid pJFP1, was used to insert uniquely labeled mutations into the DNA of *S. sanguis*.

In the work described here, four potential avirulent mutants were identified and partially characterized. To do this a number of techniques were used. One of the initial tests we did to characterize the mutants was a series of growth studies. These were performed both in the presence and absence of chloramphenicol. Initially it was decided that the growth deficit of these strains in media didn't appear to be significant enough to be the cause of the decreased signal we observed in the output pool of the initial STM experiment (Figure 3). Retesting of the putative avirulent mutants, performed by Sehmi Paik, Dr. Todd Kitten and Dr, Cindy Munro (Virginia Commonwealth University, VA) put that conclusion in doubt. In the retest, three of the four mutants (1-17, 1-36 and 1-37) produced weak signals on the inoculum blot (data not shown). We hypothesize that the decreased ability of these mutants to grow in media in the first few hours is sufficient to give them a disadvantage when co-inoculated into a rabbit with other strains.

Subsequent characterization of the mutant strains involved cloning mutant DNA fragments with the disrupted gene of interest into *E. coli*. Plasmid preps were made and the purified plasmid DNA was sequenced. Problems arose when it was determined that plasmid DNA from the vector (pJFP1), used to carry out the *in vitro* transposition reaction that created the mutants, was often inserted along with the transposon. Further analysis showed that the plasmid pJFP1 has a multitude of *Sau3AI* restriction sites (Figure 1) making it difficult to clone fragments long enough to include the entire region flanking the inserted transposon with the plasmid. Since the protocol for AP-PCR was already proving successful we decided to concentrate our efforts to capture the flanking DNA by that method rather than trying to optimize the cloning protocol.

AP-PCR worked well for characterizing the mutants but it was not without its problems. One of the first obstacles to be overcome was determining the correct concentrations of primers to use. The initial method provided by Dr. Glen Tamura (University of Washington, Seattle, Washington) was inadvertently modified to have concentrations of the arbitrary primers (Arb) greatly in excess of what had been recommended (25pmol/50  $\mu$ l reaction). This proved to be serendipitous because when the error was discovered and the protocol repeated at the recommended concentrations, it failed to work in our hands. This led us to perform a series of experiments to optimize the primer concentrations. We finally determined that 100 pmol/50  $\mu$ l reaction was the ideal concentration for the Arb primers and that the transposon-specific primers worked best at 25 pmol.

The second difficulty encountered with AP-PCR was that gels from the first round reactions would often show promise, but the gels from the second round reactions were plagued with a multitude of indistinct bands. We hypothesized that part of the problem was due to excess first round reaction primers left over in the template of the second round reactions. To "clean up" the template, we further modified the protocol to include a PCR purification step designed to remove oligonucleotides in the 17-40mer size range.

As previously mentioned in the results section, one unexpected result obtained was that identical sequences were found when trying to characterize the left end of three of the four mutants. Analysis of the M29L19 primer sequence used in the first round AP-PCR showed that it had homology at its 3' end to one of the contigs from the *S. sanguis* sequencing project. Since the first round reactions were run at a low annealing temperature, the primer was as likely to bind at that site as to anneal within the transposon. In order to overcome the problem we designed new primers. We also took into account other considerations when designing these primers. The new first round reaction primers were designed so that they were at least 90 nucleotides from the end of the transposon. The second round reaction primers were designed so that they annealed between the first primers and the inverted repeats at the edges of the transposon. Also, the second round primers were designed to be far enough apart from the first round primers such that products derived from them would be noticeably smaller than the first-round products when separated by electrophoresis on an agarose gel. Changing the primers in this manner also allowed us to eliminate any ambiguity concerning which side of the transposon gave rise to

each sequence. (Earlier experiments had employed a primer from within the transposon inverted repeat as the nested primer. Sequences obtained from this primer could be aligned by the Seqman II program with equal probability to the left or right end of the transposon.) Additionally, it ensured that the sequence was from DNA flanking the inserted transposon and not another site within the genome of *S. sanguis*.

There were two unexpected findings in the characterization of the different mutants. The first was the presence of the plasmid along with the transposon in three of the four mutants. It was expected that the exogenous transposase enzyme used in the *in vitro* transposition step of the STM experiment would cut the mini-transposon from the pJFP1 vector plasmid at both of the inverted repeat ends of the mini-transposon. Additionally, the transposase enzyme was supposed to cut the genomic DNA of *S. sanguis* anywhere there was a TA dinucleotide sequence. This would have ensured that the subsequent insertion of the transposon would essentially be random. Instead, in two of the four cases (1-39 and 1-37) the transposon was cut from the vector at only one of the two inverted repeats. This effectively linearized the plasmid, which appeared to be inserted in its entirety. In the case of 1-36, the transposase did not cut at either of the inverted repeats but instead appears to have cut at a TA sequence (or two) within the vector portion of the plasmid. This linearized fragment then appears to have been inserted into the *S. sanguis* genome.

The second unexpected finding was that at least two of the mutants (1-17 and 1-39) had two disrupted genes, one on each side of the "insert". In the case of 1-17 this "insert" was the transposon only and in 1-39 it was the transposon plus the vector. What

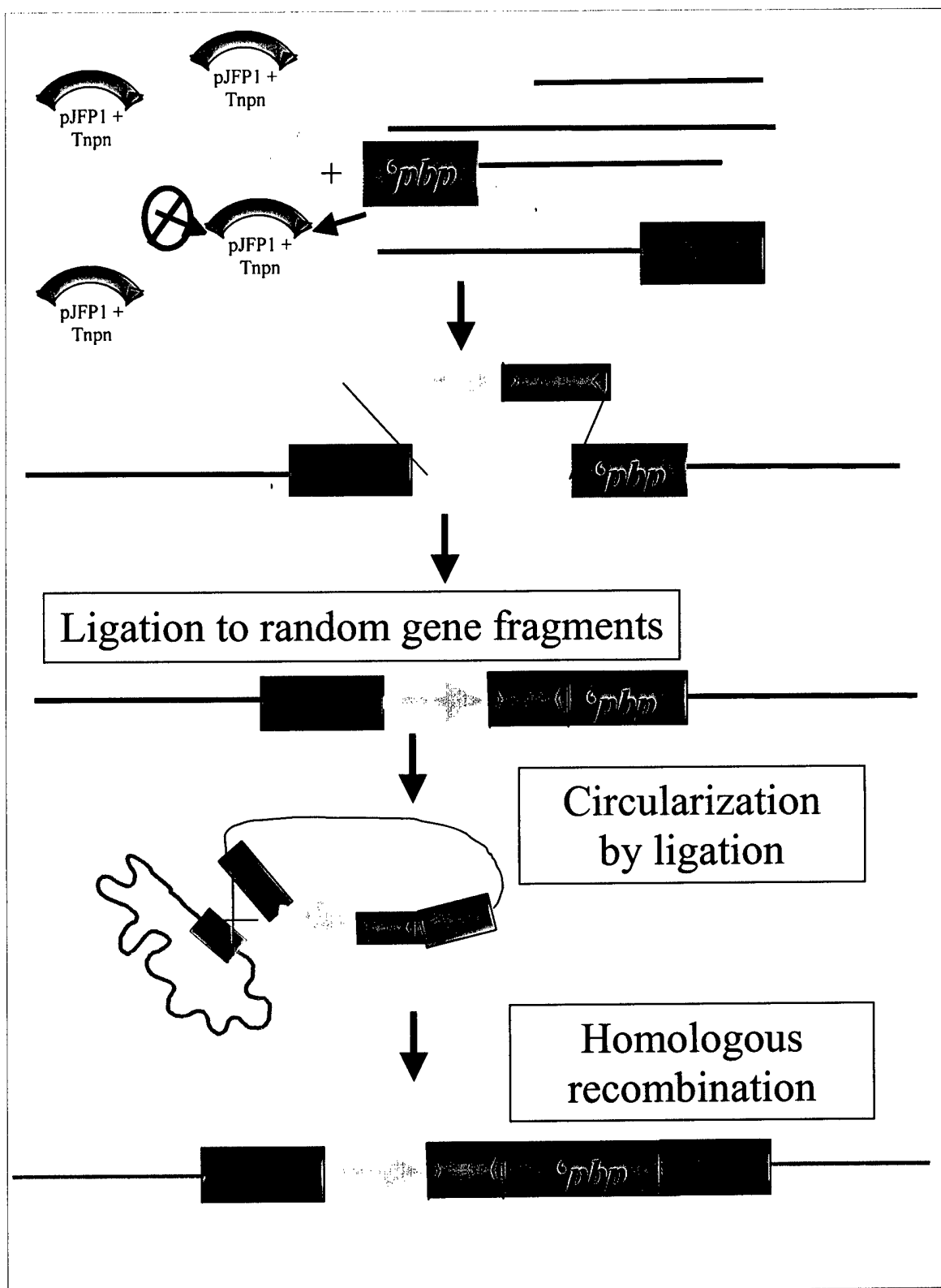
we had expected to find was a simple insertion that would have separated chromosomal sequences that were previously contiguous. This finding raised questions as to the nature of the insertion events that occurred in mutants 1-17 and 1-39. Results of the Southern blots indicated that both a wild type copy and a disrupted copy of each gene (*pgp* and *pbp-2b* in 1-39 and GTP-bp and GMP-synthase in 1-17) are present in each mutant. It also confirms that only one copy of each gene is present in the parent strain SK36.

A simple insertion model cannot explain these results. One hypothesis is that the transposon, with or without adjacent plasmid DNA, is ligated to random fragments of the *S. sanguis* genome during the in vitro ligation step of the STM procedure (Figure 15). This linear fragment could then be circularized during the ligation step, resulting in integration of the entire molecule into the *S. sanguis* genome by homologous recombination during transformation. This explanation would account for the presence of both the disrupted, non-contiguous genes flanking the transposon and the wild type copies of both genes.

Regardless of the unexpected findings encountered in trying to characterize putative avirulent genes in *S. sanguis* by using signature-tagged mutagenesis, it is still a promising method for the detection of new virulence genes. It allows for the random insertion of a uniquely "tagged" transposon into the genome of a pathogen of interest. Once potential avirulent mutants are identified they can then be effectively characterized by use of the arbitrarily primed PCR technique. In our case it appears that a problem occurred in the *in vitro* transposition step of the STM protocol. We are unsure at this time how pervasive this problem is.

Figure 15. Hypothetical Model of Mutant 1-39 Generation

This figure depicts one possible model to explain the presence of both the disrupted, non-contiguous genes flanking the transposon and the wild type copies of both genes. In this scenario the transposon is cut from the vector at only one of the two inverted repeats. The plasmid is linearized and ligated to random fragments of the *S. sanguis* genome. In the case of 1-39 the random fragments contain on one side a disrupted copy of the *pgp* gene (denoted *pgp'*) and a disrupted copy of the *pbp-2b* gene (denoted '*pbp*'). This linear fragment is then circularized during the ligation step, resulting in integration of the entire molecule into the *S. sanguis* genome by homologous recombination into the *pgp* gene during transformation. An intact copy of both genes is present. The *pgp* gene is immediately downstream of the disrupted *pbp-2b* gene and the intact *pbp-2b* gene is located somewhere else in the genome.





Studies are ongoing to see if the same types of transposition anomalies have occurred in other generations of our STM-derived mutants. Even if this is found to be the case, the overall usefulness of the technique has been proven for determining virulence factors in *S. pneumoniae* (Polissi, Pontiggia *et al.* 1998; Lau, Haataja *et al.* 2001; Hava and Camilli 2002), and we believe that our protocol can effectively be modified to preclude the anomalies from occurring in future mutants. Potential solutions include the use of less transposase enzyme or a different enzyme preparation technique, either of which might result in less digestion of chromosomal DNA, or shortening of the ligation step.

Unfortunately, none of the initial four mutants investigated will be further explored; however, the time spent trying to characterize them laid the groundwork for future analysis of other mutants and allowed us to optimize the methods that will be used.

Literature Cited

### Literature Cited

- Akerley, B. J., E. J. Rubin, et al. (1998). "Systematic identification of essential genes by in vitro mariner mutagenesis." Proc Natl Acad Sci U S A **95**(15): 8927-32.
- Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-3402.
- Baddour, L. M. (1999). "Immunization for Prevention of Infective Endocarditis." Curr Infect Dis Rep **1**(2): 126-128.
- Bayer, A. S., A. F. Bolger, et al. (1998). "Diagnosis and management of infective endocarditis and its complications." Circulation **98**(25): 2936-2948.
- Cabell, C. H., J. G. Jollis, et al. (2002). "Changing patient characteristics and the effect on mortality in endocarditis." Arch Intern Med **162**(1): 90-4.
- Dajani, A. S. (1998). "Prevention of bacterial endocarditis: Highlights of the latest recommendations by the American Heart Association." Pediatr Infect Dis J **17**(9): 824-825.
- Douglas, C. W., J. Heath, et al. (1993). "Identity of viridans streptococci isolated from cases of infective endocarditis." J Med Microbiol **39**(3): 179-82.
- Durack, D. T. (1975). "Experimental bacterial endocarditis. IV. Structure and evolution of very early lesions." J Pathol **115**(2): 81-9.
- Durack, D. T., P. B. Beeson, et al. (1973). "Experimental bacterial endocarditis. 3. Production and progress of the disease in rabbits." Br J Exp Pathol **54**(2): 142-51.
- Dyson, C., R. A. Barnes, et al. (1999). "Infective endocarditis: an epidemiological review of 128 episodes." J Infect **38**(2): 87-93.
- Guntheroth, W. G. (1984). "How important are dental procedures as a cause of infective endocarditis?" Am J Cardiol **54**(7): 797-801.

- Hanahan, D., J. Jessee, et al. (1991). "Plasmid transformation of *Escherichia coli* and other bacteria." Methods Enzymol **204**: 63-113.
- Hava, D. L. and A. Camilli (2002). "Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors." Mol Microbiol **45**(5): 1389-406.
- Hsu, S. D., J. O. Cisar, et al. (1994). "Adhesive properties of viridans streptococcal species." Microbial Ecol Health Dis **7**: 125-137.
- Kitten, T., C. L. Munro, et al. (2000). "Genetic characterization of a *Streptococcus mutans* LraI family operon and role in virulence." Infect. Immun. **68**(8): 4441-4451.
- Kitten, T., C. L. Munro, et al. (2002). "Vaccination with FimA from *Streptococcus parasanguis* protects rats from endocarditis caused by other viridans streptococci." Infect. Immun. **70**(1): 422-425.
- Kurland, S., E. Enghoff, et al. (1999). "A 10-year retrospective study of infective endocarditis at a university hospital with special regard to the timing of surgical evaluation in S-viridans endocarditis." Scand J Infect Dis **31**(1): 87-91.
- Lampe, D. J., B. J. Akerley, et al. (1999). "Hyperactive transposase mutants of the Himar1 mariner transposon." Proc Natl Acad Sci U S A **96**(20): 11428-33.
- Lampe, D. J., M. E. Churchill, et al. (1996). "A purified mariner transposase is sufficient to mediate transposition in vitro [published erratum appears in EMBO J 1997 Jul 1;16(13):4153]." Embo J **15**(19): 5470-9.
- Lampe, D. J., T. E. Grant, et al. (1998). "Factors affecting transposition of the Himar1 mariner transposon in vitro." Genetics **149**(1): 179-87.
- Lau, G. W., S. Haataja, et al. (2001). "A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence." Mol Microbiol **40**(3): 555-71.
- Maniatis, T., E. F. Fritsch, et al. (1982). Molecular cloning: a laboratory manual. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory.
- Moreillon, P., C. D. Overholser, et al. (1988). "Predictors of endocarditis in isolates from cultures of blood following dental extractions in rats with periodontal disease." J Infect Dis **157**(5): 990-5.
- Mylonakis, E. and S. B. Calderwood (2001). "Infective Endocarditis in Adults." N Engl J Med **345**(18): 1318-1330.

- Polissi, A., A. Pontiggia, et al. (1998). "Large-scale identification of virulence genes from *Streptococcus pneumoniae*." Infect Immun **66**(12): 5620-5629.
- Roberts, G. J., H. S. Holzel, et al. (1997). "Dental bacteremia in children." Pediatr Cardiol **18**(1): 24-7.
- Roberts, R. B., A. G. Krieger, et al. (1979). "Viridans streptococcal endocarditis: the role of various species, including pyridoxal-dependent streptococci." Rev Infect Dis **1**(6): 955-66.
- Sambrook, J. and D. W. Russell (2001). Molecular Cloning: a laboratory manual. Cold Spring Harbor, N.Y, Cold Spring Harbor Laboratory.
- Sullam, P. M., U. Frank, et al. (1993). "Effect of thrombocytopenia on the early course of streptococcal endocarditis." J Infect Dis **168**(4): 910-914.
- Switalski, L. M., H. Murchison, et al. (1987). "Binding of laminin to oral and endocarditis strains of viridans streptococci." J Bacteriol **169**(3): 1095-101.
- Tart, R. C. and I. van de Rijn (1991). "Analysis of adherence of *Streptococcus defectivus* and endocarditis- associated streptococci to extracellular matrix." Infect Immun **59**(3): 857-62.
- Unsworth, K. E. and D. W. Holden (2000). "Identification and analysis of bacterial virulence genes in vivo [In Process Citation]." Philos Trans R Soc Lond B Biol Sci **355**(1397): 613-22.
- Vlessis, A. A., H. Hovaguimian, et al. (1996). "Infective endocarditis: Ten-year review of medical and surgical therapy." Ann.Thorac.Surg. **61**: 1217-1222.

## VITA

Jody Christine Noe was born [redacted] in Kenosha, Wisconsin. She is a citizen of the United States of America. She graduated from Marshall High School, Marshall, Michigan in 1978. She received an Associate of Applied Science - Medical Laboratory Technician, Cum Laude, from Kellogg Community College in Battle Creek, Michigan in 1980. She then received a Bachelor of Science degree in Medical Technology from Eastern Michigan University, Ypsilanti, Michigan in 1982. As a military wife, she worked in various hospitals around the country and in Lakenheath, England as a bench level medical technologist and later as a technical supervisor until she joined the Air Force as a First Lieutenant in 1994.

Upon receiving her commission, she was assigned to the 75<sup>th</sup> Medical Group at Hill AFB, Utah initially as Associate Chief and later Chief of Clinical Laboratory Services. In 1997, she was transferred to the 366<sup>th</sup> Medical Group at Mountain Home AFB, Idaho where she was the Flight Commander of Clinical Laboratory Services. While stationed there she was selected to receive an Air Force Institute of Technology scholarship for a Master of Science in Molecular Biology.

Jody passed the American Society of Clinical Pathologists certification exams for Medical Laboratory Technician, Medical Technologist and Specialist in Hematology. She is a member of the American Society of Clinical Pathologists, the Society of Armed Forces Medical Laboratory Scientists and the American Society for Microbiology